International Journal of Women in Technical Education and Employment
ISSN: 2811-1567. Volume 5 – Issue 2. December2024
https://fpiwitedjournal.federalpolyilaro.edu.ng

Dihybrid Recurrent Neural Networks for Solar Radiation Prediction: A Comparative Study in Nigeria

Alabi, N. O. & Ojo, G.O.
Department of Mathematics & Statistics,Federal Polytechnic, Ilaro, Ogun State.
email: nurudeen.alabi@federalpolyilaro.ng.edu

Abstract

This research investigates solar radiation forecasting using two dihybrid recurrent neural networks architectures (Parallel and Sequential) on climate data from six Nigerian cities: Sokoto, Maiduguri, Ilorin, Ikeja, Enugu, and Port Harcourt. The dataset includes 31 years of monthly data on rainfall, relative humidity, sunlight hours, wind speed, maximum and lowest temperatures, and evaporation Piche. Each model is built using Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) layers and optimized through hyperparameter tuning. We analyzed solar radiation patterns, identified optimal locations for solar energy projects, and optimized agricultural planning. Both models showed excellent predictive accuracy, with the Parallel model outperforming the Sequential one. The Parallel dihybrid architecture achieved a lower Mean Square Error (MSE) of 0.0001 and a higher R^2 of 0.9995, making it more reliable for solar radiation forecasting. These findings contribute to more sustainable energy planning and agricultural optimization in Nigeria.

ARTICLE HISTORY
Received: September 23, 2024
Revised: October 30, 2024
Accepted: November 2, 2024

***Keywords**: Climate data, Dihybrid recurrent neural networks, Parallel architecture, Sequential architecture, Solar radiation forecasting, Sustainable energy planning*

Citation
Alabi, N. O. & Ojo, G.O. (2024). Dihybrid Recurrent Neural Networks for Solar Radiation Prediction: A Comparative Study in Nigeria. *International Journal of Women in Technical Education and Employment,* 5(2), 62 – 71.

Introduction

Solar radiation forecasting plays a crucial role in planning and managing renewable energy, particularly solar power. Accurate predictions can improve the efficiency of solar systems, optimize energy output, and enhance the sustainability of renewable projects. Nigeria, Africa's most populous nation, grapples with significant energy challenges, including insufficient electricity supply, high costs, and heavy reliance on fossil fuels. With a population exceeding 200 million, energy demand is set to rise, intensifying energy poverty and environmental degradation. However, Nigeria is blessed with abundant solar radiation, offering a promising renewable energy source. Regional variations in solar radiation range from 6.5 to 7.0 kWh/m² in Northern Nigeria and 5.5 to 6.0 kWh/m² in the South, as reported by the International Energy Agency (IEA, 2015). Despite a total electricity generation capacity of 12,000 MW, the country struggles to meet daily consumption needs of 30 to 40 million kWh, resulting in persistent shortages. Nigeria's reliance on fossil fuels further exacerbates climate change, air pollution, and economic instability.

Harnessing solar energy more effectively requires precise forecasting methods. This study analyses climate datasets that include solar radiation to uncover patterns and trends in Nigerian cities. Accurate solar radiation forecasting helps energy planners predict generation levels, balance supply and demand, and integrate solar power into the grid more effectively. Optimized forecasts can reduce operational costs, improve return on investment, and foster greater adoption of solar energy, thereby reducing reliance on fossil fuels and minimizing greenhouse gas emissions. Furthermore, solar radiation data is critical for agriculture, influencing crop growth, irrigation, and overall production. As Turyasingura et al. (2019) highlight, accurate climate forecasts are essential for agricultural planning, water resource management, and urban development. Recent advances in machine learning, particularly neural networks, offer enhanced accuracy and reliability in climate predictions. Neural networks, inspired by the brain's learning processes, excel in identifying patterns and generalizing across vast datasets. In climate research, they are particularly effective for predicting complex phenomena such as temperature, precipitation, solar radiation, and extreme weather events. Neural networks improve the geographic precision of predictions and capture non-linear relationships between climate variables. This study focuses on the application of neural networks to solar radiation forecasting in Nigeria, exploring innovative methodologies and challenges in this rapidly evolving field. As climate data becomes increasingly available from diverse sources like weather stations, satellites, and models, data-driven techniques such as Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU) are proving valuable in climate forecasting (Mu and Zeng, 2019). This research aims to develop a hybrid model that combines LSTM and GRU strengths, enhancing prediction accuracy and providing insights into solar radiation forecasting, data visualization, and machine learning.

Statistical analysis is indispensable in climate research, offering a powerful means to analyze climate parameters such as temperature, precipitation, and sea level pressure. By identifying patterns, trends, and uncertainties, statistical approaches allow scientists to assess climatic variability and the impacts of climate change. Importantly, statistical techniques also play a key role in evaluating climate models, which rely on complex mathematical equations to simulate Earth's climate system. As highlighted by Storch & Zwiers (2001), these techniques aid in detecting trends and understanding extreme events. Through statistical analysis, the strengths and limitations of climate models can be identified, ultimately improving predictions of climate change impacts across various sectors, including agriculture, water resources, and human health.

Research has shown that statistical methods help policymakers develop effective strategies. Studies like those of Mudelsee (2014) demonstrate the utility of time series analysis for uncovering trends in climate data. In a comprehensive assessment, Wilks (2006) focused on the utility of spatial and temporal analyses in understanding precipitation patterns. Meanwhile, Storch & Navarra (1995) explored empirical orthogonal functions (EOFs) and singular spectrum analysis (SSA) to uncover climatic patterns. These methods have proven to be vital in analyzing natural climate variability and aiding researchers in understanding long-term trends.

Artificial intelligence (AI) has emerged as another important tool in climate research, playing a significant role in addressing global climate change, as noted by Cowls et al. (2023). AI techniques such as artificial neural networks (ANN) are proving superior to traditional methods in solar radiation prediction. The accuracy of ANN models depends on input parameters, training algorithms, and architecture configurations. However, Cowls et al. (2023) also warn of AI's social, ethical, and environmental impacts, including the carbon footprint associated with AI training.

In Nigeria, solar radiation forecasting is critical for optimizing energy production and agricultural planning. Abayomi et al. (2019) developed an LSTM-based model that achieved over 97% correlation accuracy in predicting solar radiation and meteorological variables. Akpootu et al., (2019) conducted comparative studies on solar radiation estimation models in Nigerian cities, finding the quadratic exponential model effective for global solar radiation estimation. Building on these findings, this research leverages Recurrent Neural Networks (RNNs) such as Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU) to further explore solar radiation patterns in Nigeria. Recurrent Neural Networks (RNNs) are neural network architectures that have demonstrated considerable potential for modeling temporal correlations in data. RNNs are classified into two types; LSTM and GRU networks. These networks have been widely applied in various fields, including speech recognition, language modeling, and time series forecasting. Hochreiter and Schmidhuber (1997) introduced LSTM networks as an extension to regular RNNs. LSTM networks are intended to learn long-term relationships in data, solving the vanishing gradient issue that affects regular RNNs. LSTM networks are made up of memory cells, input gates, output gates, and forget gates that work together to control the flow of information. Cho et al. (2014) presented GRU networks, which are a simpler alternative to LSTM networks. GRU networks also try to learn long-term dependencies, but with fewer parameters, resulting in greater computing efficiency. GRU networks are made up of update and reset gates that govern the flow of information. Both LSTM and GRU networks are useful for modeling complex temporal relationships in data. In this study, we leverage on the capacity of these two architectures to learn patterns and trends from climate data to build a powerful dihybrid model which was used to estimate solar radiation in Nigeria. LSTM and GRU networks' functional relationship between solar radiation and other climatic factors is defined by the network's learned weights and biases.

By combining statistical analysis and machine learning, this study will enhance understanding of solar radiation variability and contribute to sustainable energy planning and agricultural strategies.

Materials and Methods

Materials

In this study, historical climate data from the Nigeria Meteorological Agency's (NiMeT) database, spanning 31 years was analyzed. The data was extracted for seven cities across Nigeria, namely Ikeja, Sokoto, Maiduguri, Enugu, Ilorin, and Port Harcourt on monthly basis. The climatic factors collected include solar radiation (*sr*) measured in (*ml*) (gunn-bellani), sunshine hours (*sh*), evaporation (*ev*) measured in (*ml*), relative humidity (*rh*) measured in percentage, minimum temperature (tmin) in degree Celsius, maximum temperature (*tmax*) in degree Celsius, rainfall (*rf*) measured in (mm) and wind speed (*ws*) measured in (m/s). The selection of cities covers different regions in Nigeria, representing various climate zones. Ikeja and Ilorin represent the western region, Sokoto from the northwest, Maiduguri from the north-eastern region, and Enugu and Port Harcourt in the southeast and south-south regions, respectively. The climatic variables were selected based on their relevance to solar radiation prediction and availability in the database. The 31-year monthly data period allows for a comprehensive analysis of climate trends and patterns in Nigeria.

Methods

Sequential and Parallel Dihybrid Models

Two dihybrid models that leverages both LSTM and GRU by combining the strengths of both architectures were trained using the climatic dataset.

(a) The Architecture of the Sequential Dihybrid Model

The Dihybrid model combines Long Short-Term Memory (LSTM) networks and Gated Recurrent Units (GRU) within a unified architecture to leverage the strengths of both in modeling temporal dependencies in climate data. LSTM networks are known for their ability to handle long-term dependencies through complex gating mechanisms, while GRUs offer a more computationally efficient structure, particularly for shorter sequences or less complex time dependencies. In the dihybrid model, the input data x_t passes through an LSTM layer which computed the hidden state h_t^{LSTM} and the cell state C_t^{LSTM} using the following equations

$$h_t^{LSTM}, C_t^{LSTM} = LSTM(x_t, h_{t-1}^{LSTM}, C_{t-1}^{LSTM}) \qquad (1)$$

In equation (1), the LSTM layer controls the flow of information through its input i_t, for gate f_t and output gate o_t which is defined as

$$i_t = \sigma(W_i x_t + U_i h_{t-1}^{LSTM} + b_i) \qquad (2)$$

$$f_t = \sigma(W_f x_t + U_f h_{t-1}^{LSTM} + b_f) \qquad (3)$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1}^{LSTM} + b_o) \qquad (4)$$

The cell state C_t^{LSTM} is updated as

$$C_t^{LSTM} = f_t \odot C_{t-1}^{LSTM} + i_t \odot \widetilde{C}_t \qquad (5)$$

Where $\widetilde{C}_t$ is the candidate cell calculated using:

$$\tilde{c}_t = ReLU(W_c \cdot x_t + U_c \cdot h_{t-1}^{LSTM} + b_c \qquad (6)$$

$$ReLU\,(x_t) = \max\,(0, x_t)$$

The output from the LSTM layer h_t^{LSTM} is fed into a GRU layer which processes it to compute the final hidden state h_t^{GRU}. The GRU layer updates the hidden state using the update gate z_t and resets gate r_t as shown below

$$z_t = \sigma(W_i x_t + U_i h_{t-1}^{GRU} + b_i) \quad (7)$$

$$r_t = \sigma(W_r x_t + U_r h_{t-1}^{GRU} + b_r)$$

$$\widetilde{h_t}^{GRU} = \tanh(W \cdot x_t + r_t \odot (U \cdot h_{t-1}^{GRU}) + b) \quad (8)$$

$$h_t^{GRU} = GRU(x_t, h_t^{LSTM}, h_{t-1}^{GRU})$$

$$= (1 - z_t) \odot h_{t-1}^{GRU} + z_t \odot Leaky\_ReLU\,(\widetilde{h_t}^{GRU}) \quad (9)$$

$$\hat{y}_t = W_y\, h_t^{GRU} + b_y \quad (10)$$

The final output of the dihybrid model h_t^{GRU}, $\hat{y}_t$ combines the long-term memory capabilities of the LSTM with the efficiency and flexibility of the GRU. This interaction allows the Dihybrid model to capture a broader range of temporal patterns, improving the accuracy of the solar radiation predictions.

(b) The Architecture of the Parallel Dihybrid Model

In a concatenated or parallel dihybrid model, LSTM and GRU networks operate on the same input data in parallel and their outputs are concatenated to produce the final prediction. The parallel networks differ in architecture each capturing different aspects of the data. We define the input sequence $[\boldsymbol{X} = X_1, X_2, \dots\dots, X_T,]$ The input x_t is fed into both the LSTM and GRU networks. The LSTM processes the input and outputs a hidden state sequence

$$[h_1^{LSTM}, h_2^{LSTM}, \dots\dots\dots, h_T^{LSTM}\ ]$$

At each time step *t*, the forget gate, input gate and output gate are

$$i_t = \sigma(W_i X_t + U_i h_{t-1}^{LSTM} + b_i) \quad (11)$$

$$f_t = \sigma(W_f X_t + U_f h_{t-1}^{LSTM} + b_f) \quad (12)$$

$$o_t = \sigma(W_o X_t + U_o h_{t-1}^{LSTM} + b_o) \quad (13)$$

$$\tilde{c}_t = tanh(W_c \cdot X_t + U_c \cdot h_{t-1}^{LSTM} + b_c \quad (14)$$

$$C_t^{LSTM} = f_t \odot c_{t-1} + i_t \odot \tilde{c_t} \quad (15)$$

$$h_t^{LSTM} = o_t \odot \tanh(c_t) \quad (16)$$

where W_*, U_*, b_* are weight matrices and bias vectors, and σ is the sigmoid function. The GRU network processes **X** and output a hidden state sequence

$$[h_1^{GRU}, h_2^{GRU}, \dots\dots\dots, h_T^{GRU}\ ]$$

At each time step *t*,

$$z_t = \sigma(W_z X_t + U_z h_{t-1}^{GRU} + b_z) \quad (17)$$

$$r_t = \sigma(W_r X_t + U_r h_{t-1}^{GRU} + b_r) \quad (18)$$

$$\widetilde{h_t}^{GRU} = \tanh(W_h \cdot X_t + U_h\, r_t \odot (h_{t-1}^{GRU}) + b_h) \quad (19)$$

$$h_t^{GRU} = (1 - z_t) \odot h_{t-1}^{GRU} + z_t \odot hard\_sigmoid\,(\widetilde{h_t}^{GRU}) \quad (20)$$

where W_*, U_*, b_* are weight matrices and bias vectors. The final hidden state from the LSTM and GRU networks are then concatenated as

$$\boldsymbol{h}^{concat} = [\boldsymbol{h}^{LSTM}, \boldsymbol{h}^{GRU}]$$

where $\boldsymbol{h}^{concat}$ is the concatenated feature vector. The concatenated vector $\boldsymbol{h}^{concat}$ passes through a dense layer to produce the final prediction

$$\hat{y} = Dense\,(\boldsymbol{h}^{concat}) \quad (21)$$

By combining the strengths of both LSTM and GRU, the models capture a broader range of patterns. Hybrid models often outperform single or standalone models in complex prediction tasks. The models adapt to various types of dependencies in the data, making them robust, and potentially improving their ability to capture long-term dependencies and efficient training dynamics.

Performance Metrics

The four performance measures used to evaluate the performance of the proposed models are Mean absolute error (MAE), Mean square error (MSE), Root mean square error (RMSE), and R^2. N is the number of times the summation iteration is executed.

$$\text{MSE} = \frac{1}{N}\sum_{1=1}^{N}(c_i - \tilde{c}_i)^2 \quad (22)$$

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}|c_i - \tilde{c}_i| \qquad (23)$$

RMSE and R^2 are defined as:

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{1=1}^{N}(c_i - \tilde{c}_i)^2} \qquad (24)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(c_i - \tilde{c}_i)^2}{\sum_{i=1}^{N}(c_i - \bar{c}_i)^2} \qquad (25)$$

c_i denotes the actual value. $\tilde{c}_i$ denotes the estimated value. N is the number of times the summation iteration is executed. These performance measures provide different insights into the accuracy and robustness of the models used for prediction. MAE measures the average magnitude of the errors in a set of predictions, without considering their direction. It is the average of the absolute differences between prediction and actual observation over the test sample where all individual differences have equal weight. MSE measures the average of the squares of the errors, which is the difference between the estimator and what is estimated. It is more sensitive to outliers than MAE.

Results and Discussion

The data on climatic factors were structured into cross-sections with each city represented by a panel. Each panel comprises 372 observations making a total of 2,232 observations available for training and testing the algorithms. Figure 1 is a visual representation of solar radiation across different periods for the six cities. Each city exhibits distinct patterns and variability in solar radiation. Understanding these patterns is crucial for applications such as solar energy planning, climate studies, and agricultural forecasting. The plot shows that solar radiation levels vary significantly across cities, with Enugu and Port Harcourt showing more variability. Maiduguri and Ilorin show more consistent levels, indicating seasonal or cyclical weather patterns.

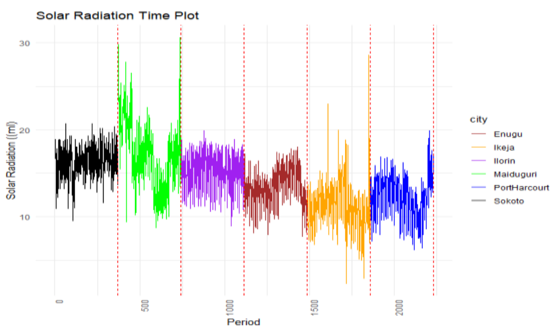

Figure 1: Time plot on Solar Radiation over Enugu, Ikeja, Ilorin, Maiduguri, Port-Harcourt, and Sokoto

Analyzing these patterns helps energy planners identify suitable locations for solar panels and optimize energy generation, with cities with higher and consistent levels being ideal candidates. Solar radiation is a key factor in agricultural productivity. Farmers and agricultural planners can use this information to make informed decisions about crop selection, planting schedules, and irrigation planning. Regions with higher solar radiation might be more suitable for crops requiring more sunlight.

Model Building and Evaluation

All of the modeling and preprocessing procedures were carried out using different libraries in *r* language. We employed *R*-Studio IDE throughout the research work. The data on the cities were arranged in the form of a panel/cross-section. This format improves the ability to analyze complex correlations, compensate for unobserved variability, and draw more robust, dynamic, and policy-relevant conclusions. The number of units, epochs, and batch sizes alongside the activation and recurrent activation functions employed in the input, forget, and output gates are shown in Table 2. The two dihybrid models were built using hyperparameter tuning to determine the best models. Firstly, best optimizer selection was conducted on some optimizers used in RNN. These are the Stochastic Gradient Descent (SGD), Adaptive Moment Estimation (ADAM), Root Mean Square Propagation (RMSPROP), Adaptive Gradient Algorithm (ADAGRAD), Adaptive Delta (ADADELTA), Nesterov-accelerated Adaptive Moment Estimation (NADAM), AdaMax (ADAMAX).

Table 1: Selection of best optimizer for Dihybrid Models

	Model	
Optimizer	LSTM	GRU
SGD	0.0669(0.2030)	0.0144(0.0958)
RMSPROP	0.0044(0.0512)	0.0033(0.0446)
ADAGRAD	0.0675(0.2120)	0.0417(0.1660)
ADADELTA	0.0093(0.0779)	0.0161(0.1080)
ADAM	0.0082(0.0734)	0.0032(0.0445)
ADAMAX	0.0209(0.1160)	0.0114(0.0844)
NADAM	0.0034(0.0462)	0.0031(0.0447)

Optimizers' MSE with MAE in parenthesis

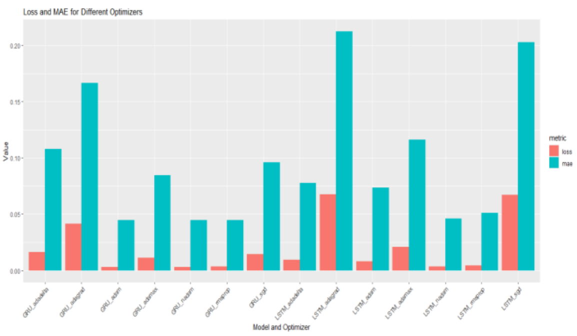

Figure 2: Bar Chart on MSE and MAE of various optimizers

Table 1 and Figure 2 show the performance of LSTM and GRU using the selected optimizers found in RNN. The table contains performance metrics such as the Means Square Error (MSE) and Mean Absolute Error (MAE). The results show that NADAM performed best amongst the optimizers. However, GRU consistently outperformed LSTM across the majority of optimizers in terms of lower MSE and MAE.

Table 2: Training Summary of Dihybrid Architectures using Hyperparameter tuning

	Parallel	**Sequential**
Optimizer	Nadam	Nadam
Unit	150	150
Batch Size	16	16

Learning rate	0.01	0.01
Activation Function	Tanh	Tanh
Recurrent Activation Function	Hard Sigmoid	Leaky ReLU
Epochs	150	150
Trainable parameters	476,851	227,251

Each of the algorithms was trained using the tuned parameters in Table 2. Both architectures use the Nadam optimizer which combines the strengths of Nesterov Accelerated Gradient (NAG) and Adaptive Moment (Adam) Estimation optimizers. This optimizer allows for the avoidance of local minima, improves convergence, and adapts the learning rate for each parameter based on the size of the gradient making it more robust and efficient. The Parallel dihybrid model has an input shape that indicates at each time step, only one feature is fed into the model, with flexible batch size. The network has two LSTM layers and two GRU layers, each with 150 units. These layers are concatenated before passing the output to a dense layer. The training process involved 10-time steps. The total number of parameters in the model is 476,851, which are all trainable. The Sequential Dihybrid model on the other hand consists of three layers. The first is an LSTM layer with 91,200 parameters, outputting a shape with unfixed batch size, sequence length (number of time steps) of 10 and 150 number of units or neurons. The second layer is a GRU layer with 135,900 parameters, outputting a shape with flexible batch size and 150 output from the GRU layer. The GRU layer takes the entire sequence from the LSTM layer and outputs just one set of 150 features, which collapses the time dimension from 10 to 1. The final layer is a fully connected layer (dense) with 151 parameters, outputting a shape such that for each sample of any batch size, the layer outputs a single value, typically a scalar prediction. Similarly, the total number of parameters in the model is 227,251 with no non-trainable parameters. In the model building process, each architecture employs 150 units per recurrent layer and is trained with a batch size of 16, a learning rate of 0.01, and for 150 epochs. However, they differ in their choice of recurrent activation function. The Parallel architecture uses Tanh as the activation function and Leaky ReLU as the recurrent activation function, while the Sequential architecture uses Tanh and Hard Sigmoid as the activation and recurrent activation functions, respectively. Leaky ReLU and Hard Sigmoid recurrent activation functions enhance performance by allowing small input fractions, gating input, learning complex patterns, and improving prediction accuracy. Equal values for units, learning rate, optimizer, activation functions, batch size, and epochs, provide a consistent basis for comparing models' performance. Residuals, seasonal, and trend patterns were included in the features of the two models. Including these components allows the models to capture complex patterns and relationships, reducing errors and overfitting thereby improving the accuracy, reliability, and robustness of predictions.

Models' Performance Comparison

The Parallel and Sequential models show excellent performance with very high R^2 values. However, the Parallel model slightly outperforms the Sequential and the standalone models in all metrics which indicates its superior accuracy and efficiency in forecasting the solar radiation dataset, despite similar RMSE and MAE values.

Table 2: Performance metrics on the Dihybrid and standalone models

Architecture	MSE	MAE	RMSE	R^2

Parallel	0.0001	0.0076	0.0076	0.9995
Sequential	0.0003	0.0122	0.0122	0.9983
GRU	0.0210	0.1088	0.1088	0.6311
LSTM	0.0382	0.1466	0.1466	0.3365

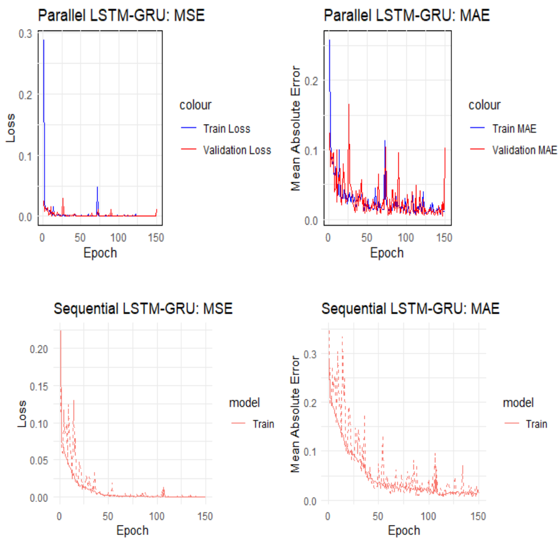

Figure 3: Training History of Parallel and Sequential Dihybrid Models

The two Dihybrid models are generally stable, with low loss (MSE) and MAE values over epochs, while there are periodic spikes (Figure 3). The Dihybrid models' continuously low MSE and MAE values indicate that they learned the data well and can generate reliable predictions. The spikes in the plots represent periods when the models struggled to match the data precisely, and were caused by noise or unexpected changes in the validation set. The Sequential model has slightly higher MSE, MAE, and RMSE values, indicating lower performance compared to the Parallel model. The Parallel model has extremely low MSE, MAE, and RMSE values, coupled with a near-perfect R^2 value, indicating an excellent fit to the data. Figure 3 shows that the rapid drop in MSE and MAE, points in the direction that our models quickly learn to reduce error during the first training steps. The stabilization of further training narrows down and indicates that errors are not significantly reduced by additional training. The convergence of training and validation MSE suggests good generalization capabilities.

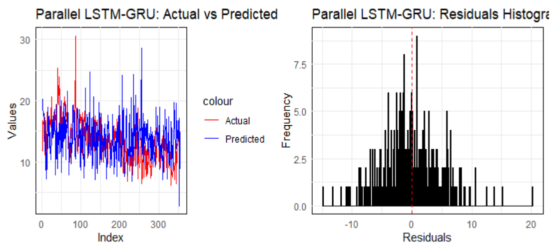

Figure 4: Actual Values versus Predicted Plot and Residual Histogram on the Parallel Dihybrid Model using the test dataset

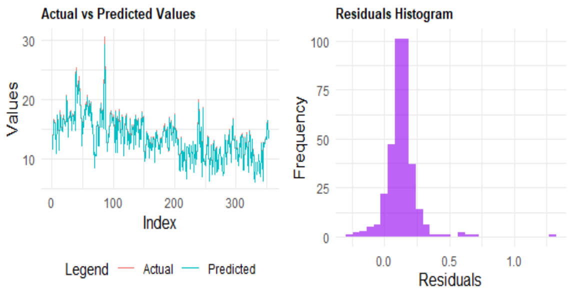

Figure 5: Actual Values versus Predicted Plot and Residual Histogram on the Sequential Dihybrid Model using the test dataset

Figures 4 and 5 present the performance of the two dihybrid models in predicting solar radiation using the test datasets. The left panels show the predicted and actual values of solar radiation, where predicted values closely track actual values for most indices, demonstrating high accuracy. Both models exhibit minimal errors, with residuals centered around zero, indicating small prediction errors. This high alignment with actual values and a narrow residual distribution highlights the models' strong forecasting capabilities. However, while the Parallel model shows a wider residual distribution, likely due to outliers, it generally performs better overall. The actual versus predicted plots suggest that the Sequential model appears more accurate visually due to tighter clustering of predictions, though this does not fully reflect the magnitude of errors. Despite this graphic difference, both models are reliable for solar radiation forecasting, with the Parallel model showing greater deviations but still excelling overall.

Conclusions

This study demonstrated the effectiveness of combining LSTM and GRU in dihybrid architectures for predicting solar radiation. By evaluating both Parallel and Sequential models, we confirmed that the Parallel architecture consistently performed better, achieving higher accuracy with lower error rates. The results support the adoption of the Parallel dihybrid model for applications requiring precise solar radiation forecasts, such as renewable energy planning and agricultural decision-making. This research highlights the potential for deep learning models to drive advancements in energy sustainability and climate resilience, particularly in regions with significant variability in solar radiation patterns.

References

Abayomi-Alli, A., Odusami, M. O., Abayomi-Alli, Misra, O. S. & Ibeh, G. F (2019). Long Short-Term Memory Model for Time Series Prediction and Forecast of Solar Radiation and Other Weather Parameters. 19th *International Conference on Computational Science and Its Applications (ICCSA)*, St. Petersburg, Russia, 82-92, Doi: 10.1109/ICCSA.2019.00004.

Akpootu, D. O., Tijjani, B. I., & Gana, U. M. (2019). Sunshine and Temperature Dependent Models for Estimating Global Solar Radiation Across the Guinea Savannah Climatic Zone of Nigeria. *American Journal of Physics and Applications* *7*(5), 125-135. Doi: 10.11648/j.ajpa.20190705.15.

Cho, K., van Merrienboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. *In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 1724–1734). Doha, Qatar: Association for Computational Linguistics. DOI: 10.3115/v1/D14-1179.

Cowls, J., Tsamados, A., Taddeo, M. et al (2023). The AI gambit: leveraging artificial intelligence to combat climate change—opportunities, challenges, and recommendations. *AI & Soc* 38, 283–307. https://doi.org/10.1007/s00146-021-01294-x.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780. DOI: 10.1162/neco.1997.9.8.1735.

IEA (2015), *Energy Statistics of OECD Countries*. OECD Publishing, Paris, https://doi.org/10.1787/energy_stats_oecd-2015-en.

Mu, R. & Zeng, X. (2019). A review of deep learning research. *KSII Transactions on Internet and Information Systems*, 13(4):1738-1764.

Mudelsee M. (2014). Trend analysis of climate time series: A review of methods, Earth-Science Reviews, 190, 310-322

Turyasingura, B., Hannington, N., Kinyi, H., Mohammed, F., Ayiga, N., Bojago, E., Benzougagh, B., Banerjee, A., & Singh, S. (2023). A Review of the Effects of Climate Change on Water Resources in Sub-Saharan Africa. African Journal of Climate Change and Resource Sustainability, *2*(1), 84-101. https://doi.org/10.37284/ajccrs.2.1.

Wilks, D. S. (2006). Statistical analysis of precipitation patterns. *Journal of Hydrology, 329*(1-2), 142-156.

Storch, H.V. & Zwiers, F.W. (2001). *Statistical Analysis in Climate Research*, Cambridge University Press, Cambridge.

Storch, H.V. & Navarra, A. (1995). Analysis of Climate Variability: Applications of Statistical Techniques. Springer-Verlag. https://doi.org/10.1007/978-3-662-03167-4.