BAJOPMAS JOURNAL 17(1): 159 – 168
Printed in Bayero University, Kano - Nigeria.

SHORT-TERM ELECTRICAL LOAD DEMAND FORECAST USING DEEP LEARNING MODELS

Sani, M. M.; *Bello, M. I.; and Ado, M.
Department of Physics, Bayero University Kano (BUK), P.M.B 3011, Kano State, Nigeria.
***Corresponding author's email:** mibello.elt@buk.edu.ng

ABSTRACT
Electricity load forecasting is essential for ensuring steady power supply and planning. In Nigeria, electricity distribution companies face persistent challenges of transformer overloading, poor voltage regulation, and frequent power interruptions, largely due to inadequate demand prediction. This research applies deep learning models to forecast short-term electricity load demand using six months of historical data obtained from an electricity distribution company. Five deep learning models viz: Bi Gated Recurrent Unit (Bi-GRU), hybrid of Convolutional Neural Network and Gated Recurrent Unit (CNN-GRU), Temporal Convolutional Network (TCN), Long-Short Term Memory (LSTM), and Gated Recurrent Unit (GRU) were trained and evaluated using Mean Absolute Error (MAE), Root Mean Square Error, and coefficient of determination (R^2) metrics. Among them, the GRU model achieved the best performance with an average MAE of 0.0292, RMSE of 0.0514, and R^2 of 0.5237. The trained model was used to forecast the next 24 hours loads, each feeder was mapped with their corresponding transformers of various stations, using the ratings of each to estimate transformer loading percentages based on the Nigerian Electricity Regulatory Commission (NERC) standard power factor of at least 0.85. Results revealed, among other transformers Gumel the TR1 was overloaded from 6 AM of 26/02/2024 to 4 AM of 27/02/2024, it has a rating of 7.5MVA, the predicted load at 6 AM was 5.7875MW which gives 90.79 overloaded, indicating potential overload risks if corrective measures are not taken. The study demonstrates the effectiveness of GRU-based forecasting in predicting short-term load demand and provides an insight to improve transformer efficiency and highlight potential overload for better distribution and system management.

Keywords: Deep Learning, Electricity Load Forecasting, GRU, Power Distribution, Transformer Loading.

1. INTRODUCTION
Electricity load forecasting is very important in power sector and has been used for over a century. It involves using mathematical methods to predict how much electricity will be needed in the future. While older statistical methods like autoregressive models have been used for this purpose, they often fail to capture the complex and non-linear nature of electricity demand, which can be influenced by things like weather, population, and the economy [1].

In recent years, Artificial Intelligence (AI), especially deep learning, has improved the accuracy of these forecasts. Models such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) have shown great results in short-term electricity load demand forecasting [2].

In Nigeria, electricity is very important for development, but the country still struggles with an unstable power supply due to limited generation capacity. Despite efforts by the government, many areas still rely on generators and small solar systems. This gap between supply and demand shows the importance of accurate load forecasting [3].

One of the major problems faced by electricity distribution companies is overload [4], which can cause low voltage, damage to equipment, and frequent blackouts. This often happens because electricity demand is not properly predicted before it occurs, and only a few studies in Nigeria combine load forecasting with practical transformer overload assessment.

This research used deep learning models to forecast short-term electricity load demand using real historical load data from an Electricity Distribution Company. The main contribution of the study has gone beyond just forecasting, it utilizes transformer loading percentage calculations to detect potential transformer overload. This will help identify which transformers may become overloaded a day before allowing electricity distribution companies to plan and respond early.

1.1 LITERATURE REVIEW

Overloading of power transformers remains one of the causes of transformer ageing and failure in power distribution systems. [5] investigated the online monitoring of overload capacity in large transformers by measuring real-time ambient and loading conditions in compliance with International Electrotechnical Commission, IEC 60354 standards [6], which provides recommendations of oil immersed power transformers, focusing on operating temperatures and their impact on insulation's thermal ageing. Their monitoring system successfully identified that a continuous overload of about 1.3 times the rated load was possible when ambient temperature fell below 20°C. However, the approach was limited to thermal estimation, it could indicate existing overload but not predict it before occurrence. Similarly, [7] developed an overload and heavy-load management platform for main transformers in Guizhou's grid. Although it enabled load risk assessment, the system can only react based predefined limit, lacking the predictive capability needed to anticipate overload conditions before they happen.

In a related study, [8] assessed the loading condition of 27 residential transformers in the Philippines and found that 22.22% were overloaded and 48.15% underloaded. Their analysis, based purely on Microsoft Excel calculations, lacked time-series modeling for analysis of future load patterns. [9] examined transformer loadings and failure rates in Onitsha's distribution network using ETAP and SPSS software. Their findings showed that overload accounted for 22.5% of transformer failures. However, the study relied heavily on static simulations and questionnaire data, rather than real-time measurements or predictive modeling. [10] applied the Auto-Regressive Integrated Moving Average (ARIMA) model to forecast 33 kV line loadings at the Ota Substation. The model produced low MAPE values for some feeders, however, it yielded negative R^2 values for others, this shows its weakness in handling nonlinear electricity consumption behavior.

These studies highlight the limitations of conventional approaches, while valuable for historical analysis, they cannot capture the nonlinear characteristics of electricity loads. Also, they do not integrate forecasting with overload assessment, that means operators only react to overload after it happens instead of preventing it.

Deep learning methods have emerged as powerful tools for capturing temporal dependencies and nonlinear patterns in energy consumption. [2] demonstrated the effectiveness of Long Short-Term Memory (LSTM) networks in short-term load forecasting, achieving a Mean Absolute Percentage Error (MAPE) of 1.65%. Despite this, the model's performance deteriorated when applied to noisy datasets, which are common in practical distribution systems.[11] enhanced LSTM performance by combining it with the XGBoost algorithm, resulting in improved accuracy. However, the hybrid model required computational resources, which limit it being feasible for real-time application.

To address LSTM's challenges, Gated Recurrent Unit (GRU) networks have been explored as an alternative. [12] introduced an attention-based GRU that achieved high accuracy under unstable weather conditions but required extensive hyper-parameter tuning which was prone to overfitting with small datasets. [1] compared GRU and LSTM for short-term demand prediction and found GRU to be faster and more accurate in short horizons, but unable to capture long-term seasonal trends due to its simpler structure.

In contrast, tree-based ensemble models such as XGBoost have proven effective in structured data forecasting. [13] optimized XGBoost for short-term load forecasting, achieving a MAPE of 2.61%, while [14] demonstrated its superior performance in capturing sudden load spikes. But, XGBoost lacks sequential learning capability, making it unsuitable for modeling temporal dependencies which is important for day ahead energy forecasting.

Hybrid models have also been used to combine the strengths of multiple algorithms. [15] developed a Bidirectional LSTM–XGBoost hybrid that accurately predicted peak loads in community energy systems but required heavy computation, which makes the system less flexible. [16] proposed a Gene Expression Programming–Adaptive Neuro-Fuzzy Inference System (GEP–ANFIS) framework that achieved low RMSE values (0.0007) but became high computational due to exponential growth in fuzzy rules as inputs increased. [17] designed an Artificial Neural Network (ANN) model that achieved a MAPE of 4.34% and correlation of 0.911; however, the model was unstable and sensitive to minor data variations. Similarly, [18] compared two Support Vector Regression (SVR) hybrid approaches for renewable energy forecasting in Nigeria and reported R^2 values above 0.99. While accurate, SVR models require kernel function selection and is less accurate when handling missing data. [19] used LSTM for long-term electricity demand forecasting, reporting MAPE values as low as 0.01% for training and 10.54% for testing. However, their model relied on monthly average data, which smooths out short-term fluctuations that are crucial for identifying daily overload conditions at transformer level.

While these deep learning and hybrid approaches have improved forecasting accuracy, they remain largely detached from practical electrical system applications such as overload prevention. Most existing models either focus on prediction accuracy without connecting results to operational parameters or are too computationally heavy for real-time deployment.

2. MATERIALS AND METHOD

This study used historical electricity load data obtained from an Electricity Distribution Company. The dataset covered a six-month period from September 2023 to February 2024 and included station names, feeders, transformer ratings, and hourly load data in megawatts (MW).

Before the analysis, the data was preprocessed to ensure quality. Non-numeric entries such as out of supply (O/S) in the load columns were removed, missing values were filled using linear interpolation, and timestamps were standardized into uniform hourly intervals. The data was then normalized using the Min–Max scaling technique which is described in equation (1) to improve training and convergence of the models by rescaling all features to a similar range between 0 and 1.

$$X_{scaled} = (X - X_{min})/(X_{max} - X_{min}) \qquad (1)$$

where X_{scaled} is the normalized load value, X_{max} and X_{min} are the maximum and minimum load values for each feeder respectively, and X is the original load value.

The prepared dataset was divided into three parts: 85% for training, 7.5% for validation, and 7.5% for testing. Five deep learning models were used and evaluated for forecasting viz: Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Bidirectional GRU (BI-GRU), Convolutional Neural Network–GRU hybrid (CNN-GRU), and Temporal Convolutional Network (TCN). Each model used data from the previous 168 hours (equivalent to seven days) to predict the next 24-hour electricity load using the test data, that is because the model does not take the 6 months data at once because the models like LSTM and GRU struggle to learn on very long sequence and can only remember a certain number of past steps.

All implementations were carried out in Python using TensorFlow and Keras libraries for deep learning. Other libraries such as Pandas and NumPy were used for data preprocessing, and Matplotlib for data visualization. Model optimization was achieved using a grid search approach which gave the best combination of 64 hidden units, 0.1 dropout rate, batch size of 32 and epoch depending on the set early dropping.

To evaluate the models, three standard forecasting metrics were applied: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the coefficient of determination (R^2) as shown in equation 2,3 and 4 respectively. The model with the best accuracy was used for final forecasting. The best model was then used to generate 24-hour-ahead electricity load predictions for each feeder using the trained model.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|x_i - \hat{x_i}| \tag{2}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{x_i})^2} \tag{3}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(x_i - \hat{x_i})^2}{\sum_{i=1}^{N}(x_i - \hat{x})^2} \tag{4}$$

where N is the total number of run-up data that have been seen, x_i and ($\hat{x_i}$) are the i-th actual and forecasted load values, respectively, and $\hat{x}$ is the mean actual value.

Finally, the predicted load values were used to compute transformer loading percentages given in equation using electrical relationship between apparent power (S), active power (P) and power factor (PF) as expressed in equation 6 and 7 below respectively.

Transformer overload can be understood using the basic power laws of electricity. The *apparent power* S handled by a transformer is given by:

$$S = V \times I \tag{5}$$

where V is the voltage and I is the current. This apparent power which is the vector magnitude of active and reactive power consists of two components: Real (active) power P, which performs useful work and Reactive power Q, which sustains the magnetic fields in inductive loads.

The relationship is expressed as:

$$S = P + jQ, \quad |S| = \sqrt{P^2 + Q^2} \tag{6}$$

The efficiency of this power transfer is determined by the power factor, which is the ratio of true power used in an AC circuit to apparent power delivered to the circuit, expressed in equation 7 below:

$$pf = \frac{P}{S} = \cos\phi \tag{7}$$

where ϕ is the phase angle between voltage and current. Overload creates condition that increase reactive power demand which lowers power factor and results in higher current in the system, which leads to increase in copper losses of the transformer because copper loss and power factor are inversely related as shown in equation (11), which causes higher operating temperature and reduced efficienciency which if not contained will result to transformer failure

$$P_{cu} = I^2 R \tag{8}$$

From equation (7) the apparent power can be expressed as:

$$S = \frac{P}{pf} \tag{9}$$

So also, from equation (5) the current I, can be expressed as:

$$I = \frac{P}{V \bullet pf} \tag{10}$$

Using Equation (8) above, the copper loss can be expressed as;

$$P_{cu} = \frac{P^2 R}{V^2 \bullet pf^2} \tag{11}$$

The transformer loading is expressed in equation (12) below;

$$\text{Loading\%} = \frac{\text{Predicted Load(MVA)}}{\text{Rated Capacity(MVA)}} \times 100 \tag{12}$$

The analysis identifies transformers operating under normal, overloaded, or underloaded conditions, providing guidance in detecting potential overload before it happens hence preventing transformer failure due to overload.

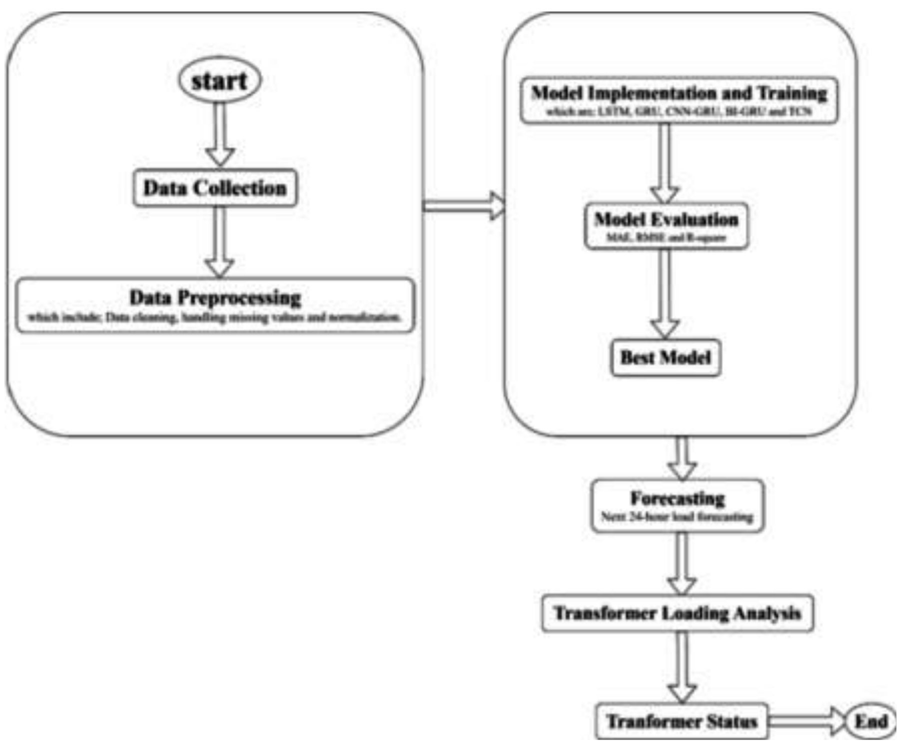


Figure 3:A flow chart showing the methodology of this research

3. RESULTS AND DISCUSSION

The validation data was used during training after each epoch to compute validation loss, validation accuracy, checks overfitting and guide early stopping. A graph of validation loss is as shown in figure 2 below. While the test data was used to evaluate the performance of the models, figure 3 shows the performance using out-of-sample forecast of Nakowa Feeder under Textile TR1 Transformer.

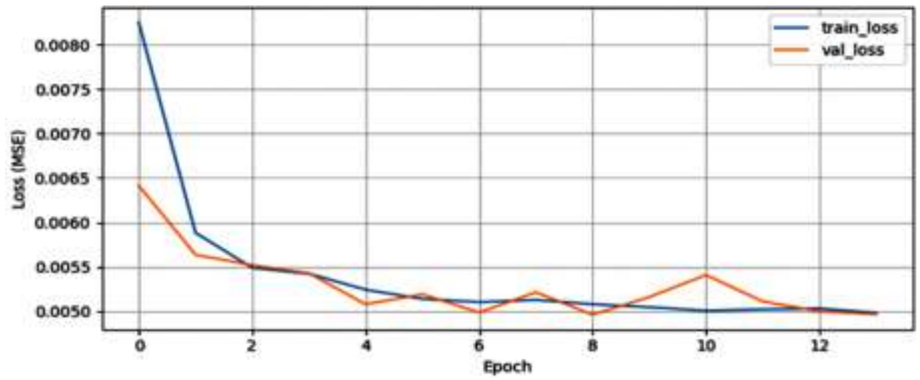
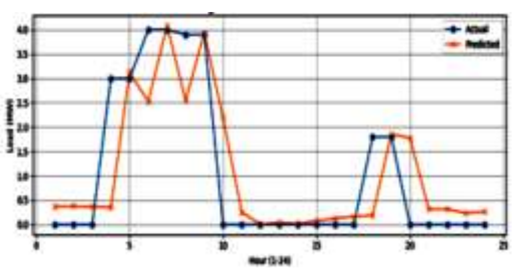
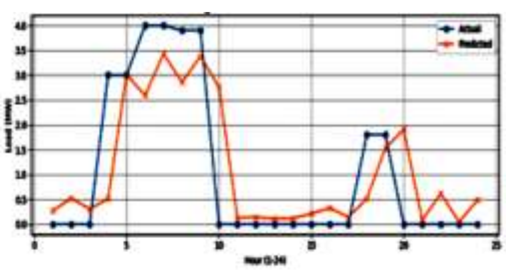


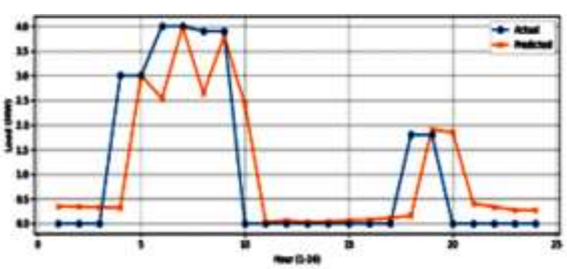
Figure 4: Validation loss curve, proving that the model is learning when the graph is descending.



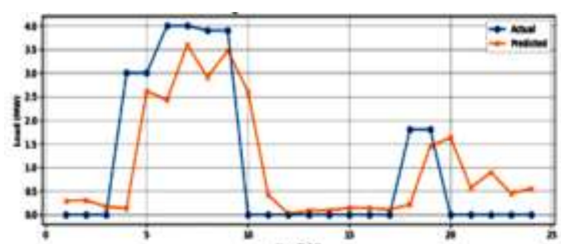
(a) BI-GRU



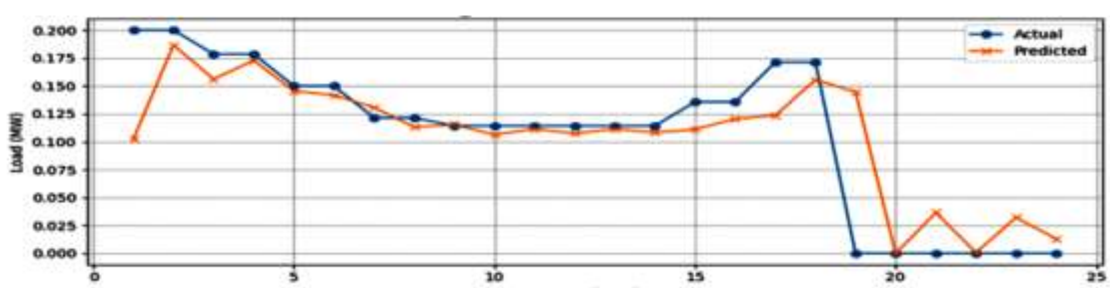
(b) CNN-GRU



(c) LSTM



(d) TCN



(e) GRU

Figure 3: the performance of the 5 deep learning Models using out-of-sample forecast.

Among all tested models CNN-GRU achieves mean absolute error (MAE) of 0.3494, root mean square error (RMSE) of 1.0883 and coefficient of determination (R^2) of 0.5330 respectively, BI-GRU achieves 0.3509, 1.1075 and 0.5308 respectively, LSTM achieves 0.3472, 1.1109 respectively and TCN achieves 0.4447, 1.1287 and 0.3370 respectively, the GRU model demonstrated the best performance, achieving the lowest mean absolute error (MAE) of 0.0292, root mean square error (RMSE) of 0.0514, and the highest coefficient of determination (R^2) of 0.5237 as shown in Table 1 below. The GRU model was used to forecast the next 24 hours of electricity load for each feeder using the trained model as shown in figure 4 below.

Table 2: Best Model based on Error Metrics

Model	Avg_MAE (MW)	Avg_RMSE (MW)	Avg_R^2
BI-GRU	0.350963	1.107577	0.530833
CNN-GRU	0.349489	1.088338	0.533062
GRU	0.029229	0.051353	0.523748
LSTM	0.347288	1.110968	0.546889
TCN	0.444709	1.128714	0.337031

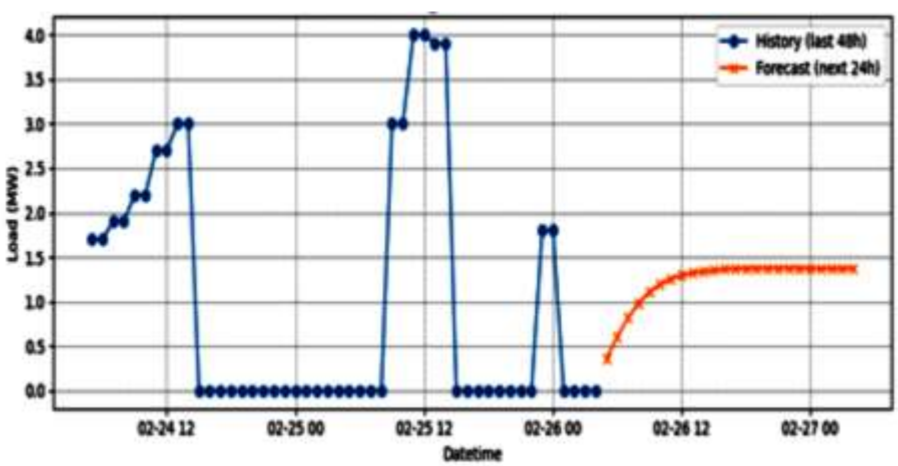


Figure 4: The next 24-hour forecast

The predicted 24-hour load values were then used to calculate transformer loading percentages for each station using a power factor of 0.85 which is the NERC standard of at least 0.85 power factor. The power factor (PF) plays a crucial role for the efficiency of the transformer. Among the station Transformers is Textile TR1 which is a 15MVA transformer that supply electricity to the following feeders: Nakowa, Mairuwa, Funtua Textile Mill and Funtua Water Works.

To evaluate the condition of each transformer, the loading percentage was computed, each station transformer was assessed hourly for 24 hours, taking Textile TR1 as an example, the predicted values of all the feeder were summed, the value obtained which is in MW was converted to MVA using equation (7) above, and hence using equation (12) above to compute the loading percentage of that transformer. The total result shows 43 normal conditions, 23 overloaded conditions while 1494 instances were underloaded as shown in Figure 5 below:

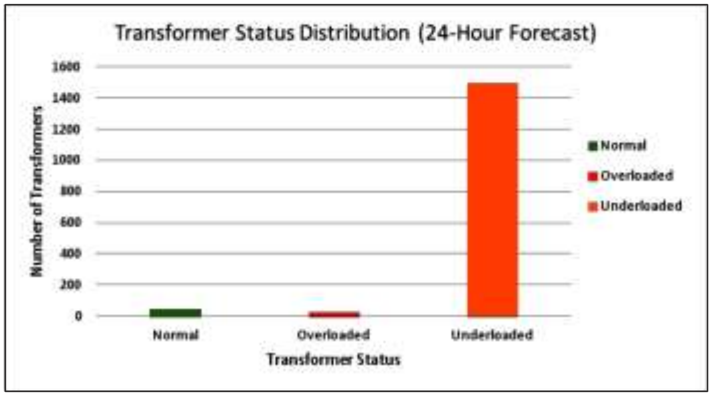


Figure 5: Transformer status distribution at different instances

The results demonstrated that the GRU model provided the most reliable short-term forecasting performance among all tested architectures. The transformer loading analysis revealed that most transformers operated under safe loading conditions, while Gumel TR1 transformer was overloaded from 6 AM of 26/02/2024 to 4 AM of 27/02/2024. Figure 6 below, shows how power factor affects the transformer performance which reduces as a result of transformer overload. Five power factors were used to compute the transformer loading percentage viz; 0.5, 0.67, 0.78, 0.8 and finally 0.85 which is the standard, the result shows that, A higher power factor reduces the chance of transformer overloading.

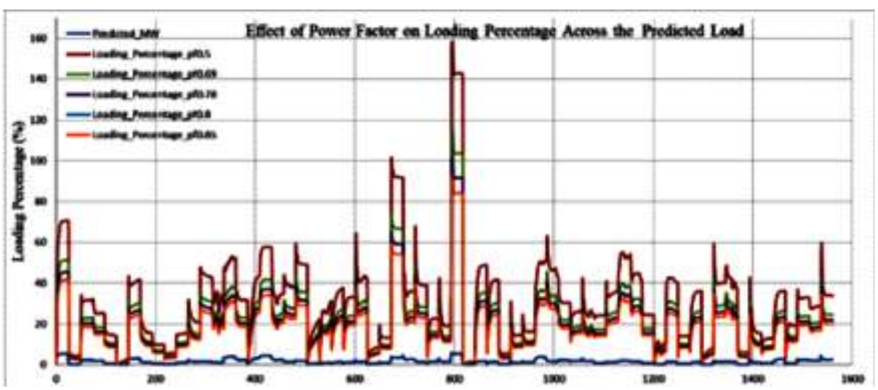


Figure 6: Effect of power factor on transformer loading percentage across the predicted load.

The findings align with those of [5] and [8], who emphasized that transformer overloads often stem from load mismanagement rather than capacity limits. Unlike earlier threshold-based systems [7], this approach allows operators to anticipate overloads before they occur. The integration of forecasting with electrical analysis represents a practical step beyond previous deep learning studies [2], [12], [14], which primarily focused on prediction accuracy without linking results to operational decisions.

4. CONCLUSION

This research demonstrated the use of deep learning for short-term load forecasting and overload risk assessment across 11 kV feeders. The study provides a framework for improving the distribution network and planning transformer maintenance in advance.

Future studies could extend this work by incorporating additional factors such as weather. Moreover, expanding the forecasting horizon to cover seasonal variations may further strengthen policy-level planning for energy distribution in Nigeria.

REFERENCES

[1] W. Wu, W. Liao, J. Miao, and G. Du, "Using gated recurrent unit network to forecast short-term load considering impact of electricity price," in *Energy Procedia*, Elsevier Ltd, 2019, pp. 3369–3374. doi: 10.1016/j.egypro.2019.01.950.

[2] J. S. Caicedo-Vivas and W. Alfonso-Morales, "Short-Term Load Forecasting Using an LSTM Neural Network for a Grid Operator," *Energies*, vol. 16, no. 23, 2023, doi: 10.3390/en16237878.

[3] T. . Maku, M. U. Adehi, and M. O. Adenomon, "Modeling and forecasting electricity consumption in Nigeria using Arima and Arimax time series models," *Sci. World J.*, vol. 18, no. 3, pp. 414–421, Oct. 2023, doi: 10.4314/swj.v18i3.14.

[4] H. Ma, P. Yang, F. Wang, X. Wang, D. Yang, and B. Feng, "Short-Term Heavy Overload Forecasting of Public Transformers Based on Combined LSTM-XGBoost Model," *Energies*, vol. 16, no. 3, 2023, doi: 10.3390/en16031507.

[5] S. Tenbohlen, T. Stirl, and M. Stach, "Assessment of overload capacity of power transformers by on-line monitoring systems," *Proc. IEEE Power Eng. Soc. Transm. Distrib. Conf.*, vol. 1, no. WINTER MEETING, pp. 329–334, 2001, doi: 10.1109/pesw.2001.917057.

[6] T. Sybel, *Transformer monitoring according to IEC 60354*, Edition 06., vol. 105, no. 2006. 2007. [Online]. Available: www.a-eberle.de

[7] F. Qiang *et al.*, "Application of transformer heavy load and overload analysis and management system," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 657, no. 1, 2021, doi: 10.1088/1755-1315/657/1/012097.

[8] R. H. Gomez, J. L. M. Dizon, C. A. S. Teodoro, J. C. Demegillo, and N. T. Florencondia, "Electrical Percent Loading Assessment for the Distribution Transformers Residential-Used of a Barangay," *Int. J. Eng. Adv. Technol.*, vol. 9, no. 3, pp. 3020–3024, 2020, doi: 10.35940/ijeat.c5920.029320.

[9] H. N. Amadi and F. I. Izuegbunam, "Analysis of Transformer Loadings and Failure Rate in Onitsha Electricity Distribution Network," *Am. J. Electr. Electron. Eng.*, vol. 4, no. 6, pp. 157–163, 2016, doi: 10.12691/ajeee-4-6-2.

[10] F. E. Shaibu, E. N. Onwuka, N. Salawu, and S. S. Oyewobi, "Medium Term Load Forecasting Of 33 Kv Line Loading: A Case Study Of Ota 132/33 Kv Sub-Station Ogun State, Nigeria.," *Niger. J. Technol.*, vol. 43, no. 4, pp. 743–753, Feb. 2025, doi: 10.4314/njt.v43i4.14.

[11] R. A. Abbasi, N. Javaid, M. N. J. Ghuman, Z. A. Khan, S. Ur Rehman, and Amanullah, "Short Term Load Forecasting Using XGBoost," *Springer, Cham*, pp. 1120–1131, 2019, doi: 10.1007/978-3-030-15035-8_108.

[12] S. Jung, J. Moon, S. Park, and E. Hwang, "An Attention-Based Multilayer GRU Model for Multistep-Ahead Short-Term Load Forecasting," *Sensors*, vol. 21, no. 5, 2021, doi: 10.3390/s21051639.

[13] Z. Qinghe, X. Wen, H. Boyan, W. Jong, and F. Junlong, "Optimised extreme gradient boosting model for short term electric load demand forecasting of regional grid system," *Sci. Rep.*, vol. 12, no. 1, p. 19282, 2022, doi: 10.1038/s41598-022-22024-3.

[14] M. Mohamed, F. E. Mahmood, M. A. Abd, M. Rezkallah, A. Hamadi, and A. Chandra, "Load Demand Forecasting in Smart Microgrid Using NWP via Feature Engineering," in *2023 IEEE International Conference on Energy Technologies for Future Grids (ETFG)*, Wollongong, Australia: IEEE, Dec. 2023, pp. 1–6. doi: 10.1109/ETFG55873.2023.10408549.

[15] L. Semmelmann, S. Henni, and C. Weinhardt, "Load forecasting for energy communities: a novel LSTM-XGBoost hybrid model based on smart meter data," *Energy Informatics*, vol. 5, Sep. 2022, doi: 10.1186/s42162-022-00212-9.

[16] M. S. Bakare, A. Abdulkarim, A. N. Shuaibu, and M. M. Muhamad, "A hybrid long-term industrial electrical load forecasting model using optimized ANFIS with gene expression programming," *Energy Reports*, vol. 11, pp. 5831–5844, Jun. 2024, doi: 10.1016/j.egyr.2024.05.045.

[17] C. U. Eya, A. O. Salau, S. L. Braide, and O. D. Chigozirim, "Improved Medium Term Approach for Load Forecasting of Nigerian Electricity Network Using Artificial Neuro-Fuzzy Inference System: A Case Study of University of Nigeria, Nsukka," in *Procedia Computer Science*, Elsevier B.V., Jan. 2023, pp. 2585–2593. doi: 10.1016/j.procs.2023.01.232.

[18] S. I. Abba *et al.*, "Emerging Harris Hawks Optimization based load demand forecasting and optimal sizing of stand-alone hybrid renewable energy systems– A case study of Kano and Abuja, Nigeria," *Results Eng.*, vol. 12, Dec. 2021, doi: 10.1016/j.rineng.2021.100260.

[19] R. Bareth, A. Yadav, S. Gupta, and M. Pazoki, "Daily average load demand forecasting using LSTM model based on historical load trends," *IET Gener. Transm. Distrib.*, vol. 18, no. 5, pp. 952–962, Mar. 2024, doi: 10.1049/gtd2.13132.