FEDERAL UNIVERSITY OF TECHNOLOGY, AKURE
JOURNAL OF ENGINEERING AND ENGINEERING TECHNOLOGY

A COMPARATIVE ANALYSIS OF PREDICTIVE MODELS FOR HOUSEHOLD ENERGY CONSUMPTION IN LAGOS

[1]Raheem, W. A., [2]Omiyale, A. D., [1]Odeyinka, O. F.* and [1]Folorunso, C.

[1]Department of Systems Engineering, University of Lagos, Lagos State, Nigeria
[2]Department of Computer Science Education, Corona College of Education, Lagos, Lagos State, Nigeria

**Correspondence author: oodeyinka@unilag.edu.ng*

Raheem, W. A., Omiyale, A. D., Odeyinka, O. F. and Folorunso, C. (2025): A Comparative Analysis of Predictive Models for Household Energy Consumption in Lagos. *FUTA Journal of Engineering and Engineering Technology /19(2), 63-73.*

Received Date: 25.09.25 **Accepted Date: 17.10.25**

Abstract
Rapid urbanisation, unreliable electricity supply, and increasing reliance on electrical appliances have heightened residential energy demand and created the need for accurate prediction of household energy usage. While deep learning and hybrid predictive models have been increasingly applied to energy forecasting globally, their use in African contexts remains comparatively limited, with relatively few studies addressing the added complexities of infrastructural fragility and diverse user behaviours. In Nigeria, there has been limited effort in comparative modelling research, particularly at the household level, where energy behaviour is shaped by a complex interplay of socio-economic and infrastructural factors. This study examines household energy consumption patterns in Lagos, Nigeria, using a comparative, multi-model approach. Data from 350 households, including socio-economic status, appliance usage, billing methods, and conservation practices were collected and analyzed. The results of four predictive models, Linear Regression, Artificial Neural Networks (ANN), AutoRegressive Integrated Moving Average (ARIMA), and a hybrid ARIMA–Long Short-Term Memory (LSTM) model, were compared. Findings indicate that household income, dwelling type, and appliance usage are significant predictors of electricity consumption. Among the models, the hybrid ARIMA–LSTM achieved the best performance (RMSE of 580, MAE of 400, R^2 of 0.85), combining the strengths of both linear and non-linear modeling. The ANN model also demonstrated strong predictive accuracy (R^2 of 0.79), while linear regression was limited by its inability to capture complex relationships. The results highlight the potential of integrating machine learning into urban energy planning. Policymakers could use these insights to support targeted interventions such as smart metering, appliance subsidies, and behaviour-based demand management.

Keywords: *Artificial neural networks, autoregressive integrated moving average, energy consumption, hybrid models, linear regression, machine learning.*

Introduction

Household energy consumption constitutes a critical component of national energy planning and sustainability strategies, especially in developing nations such as Nigeria. In an urban centre like Lagos, which is Nigeria's most populous and economically vibrant city, rapid demographic expansion and industrialization have increased the ownership of home appliances (Adamu *et al*., 2020). These factors reflect socio-economic progress but have led to a corresponding surge in electricity demand (Somoye, 2023). Unfortunately, this demand exists in a national energy system plagued by instability, insufficient generation capacity, and distributional inefficiencies *(Ubani et al*., 2024).

Despite Nigeria's status as fourteenth largest oil producer in the world (Trading Economics, 2025), the country's energy infrastructure remains underdeveloped and unreliable. The power sector has been characterized by erratic supply, low per capita electricity access, frequent blackouts, and high dependence on fossil-fuel-based backup sources such as petrol and diesel generators (Onatunji, 2025). These challenges typify Lagos, the commercial nerve center of the country with more than fifty percent of functional manufacturing companies in Nigeria (Ubani et al., 2024). Urban households in Lagos are often forced to alternate between grid power and generators, leading to significant cost burdens and environmental pollution (Somoye, 2023). It has been estimated that over 85 million Nigerians are not connected to the national grid, with many more experiencing under-electrification (Ugbehe *et al*., 2025). The inefficient nature of the electricity distribution system in the country exacerbates these challenges. Many households still lack access to prepaid meters and are instead subjected to estimated billing practices, which often result in inflated and unpredictable electricity charges (Onyenandu *et al*., 2025). These systemic issues increase the cost of living and discourage proactive energy-saving behaviours (Bamidele *et al*., 2020) and contribute to a lack of accountability in energy consumption, undermining efforts to implement sustainable energy management policies (Ugbehe *et al*., 2025). With the steady rise in energy demand and its environmental and economic implications, there is an urgent need for data-driven strategies that can forecast consumption trends and inform targeted interventions. Predictive modeling presents a viable solution by enabling stakeholders to anticipate demand patterns, evaluate usage behaviours, and identify opportunities for conservation. These models, when grounded in robust socio-demographic and appliance usage data, can significantly improve the accuracy of electricity demand forecasting and support smarter infrastructure planning.

Numerous studies have demonstrated the effectiveness of hybrid and machine learning models in energy forecasting. For example, (Suhartono et al., 2012) developed a seasonal ARIMA-ANFIS hybrid model using electricity load data from Indonesia. Their approach combined the linear strengths of ARIMA with the nonlinear adaptive capabilities of Adaptive Neuro-Fuzzy Inference System (ANFIS), resulting in improved accuracy for short term load forecasts. However, the study was limited by a relatively small dataset and its focus on short range prediction. In a related study, Chaturvedi *et al.* (2022) compared four models - Seasonal Autoregressive Integrated Moving Average (SARIMA), Long Short-Term Memory (LSTM), Recurrent Neural Network (RNN), and Facebook Prophet using ten years of national energy demand data from India. Their results showed that the deep learning models, particularly LSTM and RNN, consistently outperformed traditional time series models, although they required considerable computational resources and extensive data preparation. (Giacomazzi et al., 2023) used the Temporal Fusion Transformer (TFT), (which is a deep learning framework) for multivariate forecasting, using substation level data. The model achieved a mean absolute percentage error (MAPE) of 2.52 per cent, showing strong predictive accuracy, though its complexity made interpretation more difficult. Similarly, Huang *et al*. (2022) proposed a CEEMDAN Sample Entropy BPNN Transformer hybrid model for short term load forecasting in urban China. Their model performed well in managing noisy and irregular datasets, a challenge also observed in dense cities such as Lagos, but required careful parameter adjustment to remain stable. Other deep learning-based frameworks for energy prediction have also been reported by Bourdeau *et al*. (2019), Han and Fan (2024), and Doğan *et al*. (2024). These studies confirm the potential of hybrid approaches while acknowledging continuing challenges such as data limitations, high computational demands, and limited generalization across regions.

Despite the widespread success of deep learning and hybrid predictive models in forecasting energy usage across developed and emerging economies, these methodologies have not been widely applied to the African context, where infrastructural fragility and user behaviour introduce additional complexity. In Nigeria, there has not been enough effort in comparative modelling literature, especially at the household level where energy behaviour is influenced by a complex mix of socio-economic and infrastructural variables (Adamu *et al*., 2020; Ubani *et al*., 2024). Limited researches have explored the comparative effectiveness of different predictive models within the context of residential energy consumption in Nigeria. Most studies have either focused on macro-level national forecasts or employed single-model methods without investigating the relative strengths and weaknesses of alternative approaches. Some focused on national-level demand projections or macroeconomic energy models but rarely address the nuanced, localized dynamics found in urban residential settings. As highlighted by (Nti *et al*., 2020), overlooking micro-level energy determinants such as appliance ownership, income variability, and behavioural patterns can also lead to misleading forecasts and suboptimal policy responses. Lack of access to prepaid meters and inconsistent billing systems have fragmented urban areas, resulting to unreliable datasets and ineffective policies (Bamidele *et al*., 2020). Hence, this gap requires a grounded, multi-model approach that not only forecasts consumption but also offers interpretability and contextual relevance. By incorporating features such as income, dwelling type, and appliance inventory into model design, this study seeks to produce more actionable insights for both utility managers and policymakers. This helps to bridge the methodological divide between global advances in energy analytics and the underexplored realities of residential energy use in Nigeria. Another concern is the human dimension of energy use. Research has shown that income level, education, and environmental awareness significantly influence household energy behaviour (Nti *et al*., 2020).

Hence, this study compares multi-model approach for forecasting household energy consumption in Lagos combining linear regression, artificial neural network (ANN), and AutoRegressive Integrated Moving Average into a single framework. By grounding these models in local data and combining technical evaluation with behavioural insights, this work contributes to urban energy management in Africa. The findings help to improve forecasting and guide policy instruments developed for smart metering, appliance subsidies, and behaviour-based demand management. This study also analyses behavioural patterns such as the adoption of energy-efficient appliances and conservation practices to determine their correlation with socio-economic variables. This approach aids the identification of high-impact leverage points for policy, such as subsidizing efficient appliances or promoting awareness campaigns targeted at specific income brackets. To ensure that modeling insights translate into actionable policy recommendations, our study also considers implementation feasibility, interpretability of model outputs, and alignment with Nigeria's broader energy goals. The inclusion of ARIMA which is relatively transparent and easy to operationalize, alongside more complex ANNs, offers a balance between accuracy and usability. Additionally, the SHapley Additive exPlanations (SHAP) method is incorporated to enhance the interpretability of our machine learning models and support policy transparency.

The remainder of this paper is organised as follows. Section 2 outlines the materials and methods, including data collection, preprocessing, feature development, and model construction. Section 3 presents and discusses the results, comparing the performance of the predictive models. Section 4 concludes the study with key insights and recommendations, while the final section discusses the policy implications of the findings and their relevance to urban energy planning.

Materials and Method

Error! Reference source not found. describes the model development process used in this study. It begins with data collection through a structured household survey, followed by preprocessing steps such as encoding and normalization.

The predictive models (Linear Regression, Artificial Neural Networks, and AutoRegressive Integrated Moving Average) were implemented using appropriate tools. These models were trained and evaluated based on standard metrics including R^2, MAE, MSE, and RMSE which are discussed later in this section. The final step involved comparing model performance to identify the most accurate and policy-relevant approach for forecasting household energy consumption in Lagos and its environs.

Data Collection

This study employed a cross-sectional survey distributed online to a targeted sample of households

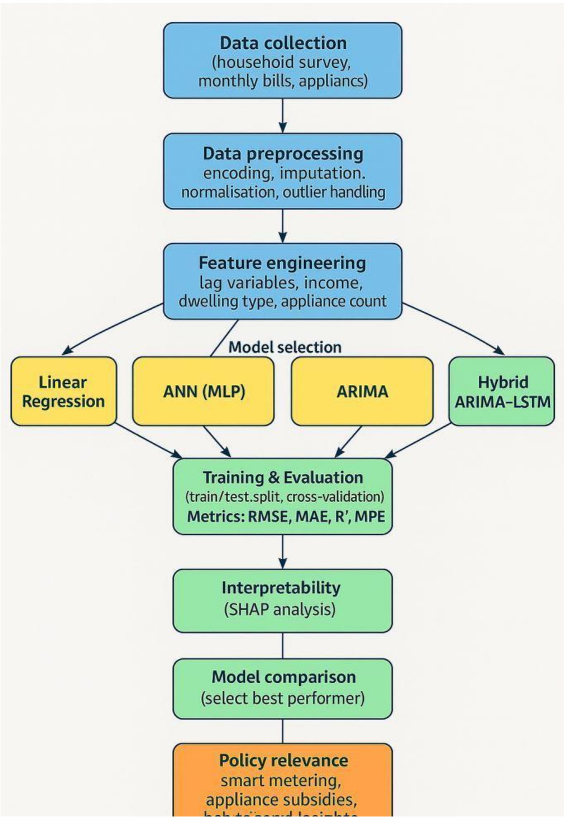

Figure 1: The Model Development Process

in Lagos, Nigeria. A stratified random sampling method was used to ensure diversity in gender, age, household income range, and housing type. The questionnaire consisted of two parts. The first section gathered demographic information, while the second was divided into four themes: household characteristics, energy consumption patterns, appliance usage, and awareness and application of energy-saving practices. Prior to distribution, the questionnaire was reviewed by two energy research experts and piloted with twenty households to assess clarity, relevance, and internal consistency. Feedback from the pilot was used to refine ambiguous questions and improve structure. Data validation was carried out through consistency checks and cross-verification of responses, such as comparing reported appliance ownership with self-reported electricity expenditure ranges. Incomplete or contradictory responses were removed from the dataset before analysis. In total, 350 valid and verified responses were retained and collated in a spreadsheet format for subsequent processing and modelling.

Data Preparation and Pre-processing

The responses were compiled and structured into a data frame in Python. Categorical data (e.g., dwelling type, meter type) were encoded numerically to enable analysis. Missing values were handled through imputation where appropriate. Outliers were identified using Z-score methods and winsorised where necessary to maintain data integrity.

Predictive Modelling Techniques

The predictive models considered in this study are discussed as follows.

Linear Regression Model

This model was developed using scikit-learn in Python. Independent variables included household income, number of appliances, meter type, and daily appliance usage. While straightforward and interpretable, the model assumes linear relationships between predictors and target variables.

Artificial Neural Network (ANN)

A feedforward Multi-Layer Perceptron (MLP) with one hidden layer of 20 neurons and ReLU activation was trained using TensorFlow and Keras. Input features were standardized, and the model used the Adam optimizer. The model was trained for 100 epochs with early stopping based on validation loss. A batch size of 32 was used, and the dataset was split into 80% training and 20% testing sets. The mean squared error was used as the loss function.

Additionally, the SHAP was applied to interpret the ANN and hybrid model outputs, to quantify the contribution of each feature to energy prediction. SHAP is a widely used method for explaining complex predictive models (Lundberg & Lee, 2017) and has been successfully applied in energy forecasting contexts (Baur *et al*., 2024).

AutoRegressive Integrated Moving Average (ARIMA)

Monthly electricity bill data covering three years (from January 2022 to December 2024) were used to develop the ARIMA model. The optimal parameters (p, d, q) were selected based on the Akaike Information Criterion (AIC), and the model was implemented in Python using the *statsmodels* library. The ARIMA approach is effective for identifying linear temporal patterns and seasonal fluctuations in energy consumption. To allow fair comparison across models, lagged variables such as the previous month's consumption were also included in the ANN and Linear Regression models.

Hybrid ARIMA–LSTM Mode

Recent advancements in energy forecasting have shown that combining statistical and deep learning models can significantly improve predictive accuracy, particularly where data exhibit both linear and nonlinear characteristics. For instance, Ruelens *et al.* (2022) demonstrated that a hybrid LSTM–XGBoost model trained on smart meter data outperformed standalone LSTM and other baseline approaches. Traditional time-series models like ARIMA (AutoRegressive Integrated Moving Average) are effective at capturing linear trends and seasonality in energy data, but they often fall short in modeling complex, non-linear dynamics that are increasingly prevalent in residential electricity consumption. Conversely, Long Short-Term Memory (LSTM) networks is a type of recurrent neural network that learns non-linear temporal dependencies, making it ideal for capturing fluctuations influenced by human behaviour, weather variation, or policy interventions (Jang *et al*., 2024; Lim *et al*., 2021).

The hybrid ARIMA–LSTM model thus integrates the strengths of both approaches. ARIMA is first used to identify and model the linear components of the time series data, effectively removing predictable trends and seasonality. The residuals representing non-linear and irregular variations unexplained by ARIMA are then passed to an LSTM network for deeper learning. In this study, the LSTM was configured with two hidden layers containing 50 units each, and a dropout rate of 0.2 to prevent overfitting. The final forecast is generated by combining the ARIMA-derived linear forecast with the LSTM-derived prediction of the residuals.

This decomposition-based hybrid architecture has been validated in various domains, including urban energy forecasting, due to its robustness and adaptability (Ji *et al*., 2022; Sharma *et al*., 2024). Particularly in contexts like Lagos, where energy consumption patterns reflect a mix of structured (billing cycles) and unstructured (generator usage) dynamics, this approach offers a balanced and highly accurate prediction method. It bridges the gap between interpretability and modelling power, which are two features often traded off in energy analytics.

Evaluation Metrics

The models were evaluated using four standard metrics commonly applied in time series and regression analysis.

a) Mean Absolute Error (MAE):

MAE measures the average size of the prediction errors, regardless of direction. It is given by Equation (1).

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | y_{1} - \hat{y}_i | \quad (1)$$

Where y_1 represents the observed value, $\hat{y}_i$ is the predicted value, and n is the total number of observations. Lower MAE values indicate greater predictive accuracy.

b) Mean Squared Error (MSE):

MSE represents the average of the squared differences between the observed and predicted values. It penalises larger errors more heavily and is expressed as shown in Equation (2):

$$MSE = \frac{1}{n}\sum_{i=1}^{n} (y_{1} - \hat{y}_i)^2 \quad (2)$$

c) Root Mean Square Error (RMSE):

RMSE is the square root of MSE and expresses the prediction error in the same units as the target variable (for example, kilowatt-hours or naira). It is calculated as shown in Equation (3).

$$RMSE = \sqrt{\frac{1}{n}\left(\sum_{i=1}^{n} (y_{1} - \hat{y}_i)^2\right)} \quad (3)$$

d) Coefficient of Determination (R²):

R² indicates how well the independent variables explain the variation in the dependent variable. It is as shown in Equation (4).

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_{1} - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_{1} - \bar{y})^2} \quad (4)$$

where $\bar{y}$ is the mean of the observed values. An R^2 value closer to 1 indicates a better model fit, meaning that most of the variation in the data is explained by the model.

These metrics together provide a balanced assessment of model performance, reflecting both the size and consistency of prediction errors.

Results and Discussion

Descriptive Analysis of Respondent Data

Table 1 summarises the key demographic characteristics of these respondents. There were 54% male respondents, while 46% were female, while 4% preferred not to disclose their gender.

Majority of respondents (64%) were between the ages of 18 and 30 years, 30% were aged between 31 and 45 years while the remaining 6% were between 46 and 60 years. This distribution reflects a relatively young respondent population, which is typical of an urban city like Lagos. For the level of income, most of the respondents earn within ₦200,001 to ₦500,000, followed by those earning more than ₦500,000. Lower-income groups, particularly those earning less than 100,000, are underrepresented in the sample. Overall, a significant portion of respondents earn above 200,000 monthly, indicating that the group surveyed leans towards a relatively wealthy demographic. These demographics provide essential context for interpreting energy usage patterns and for the development of predictive models tailored to urban household settings. **Error! Reference source not found.** illustrates key demographics of the respondents.

The majority of respondents live in one-room or self-contained units and 2 to 3-bedroom flats, typically accommodating small households of one to two people. Duplexes are the least common housing type, representing just over 3% across all household sizes. For smaller households, one-room or self-contained apartments and tiny flats are the preferred types of housing. Interestingly, larger families with five or more members tend to reside more often in one-room/self-contained units and micro apartments rather than duplexes or larger flats. This pattern suggests that compact living spaces are favored by households regardless of size, with duplexes consistently being the least prevalent choice.

Error! Reference source not found. presents the distribution of appliance and equipment inventory among respondents, categorized by both type of dwelling and household size. The dwellings include a one-room apartment, mini-flat, duplex, 2-bedroom flat, and 3-bedroom flat. The data reveals that 1–2 person households in one room apartment and 3–4 person households in 3-Bedroom Flats recorded the highest inventory levels at 12% each. Generally, larger households in larger dwellings (e.g., 3-Bedroom Flats) tend to have higher inventories, whereas Duplex dwellers across all household sizes show comparatively lower percentages, indicating possibly fewer respondents or minimal appliance accumulation.

Figure 4 illustrates the distribution of monthly household income among the respondents. The

Table 1: Demographic Summary of Respondents

Demographic Category		Subcategory	Count	Frequency (%)
Sex		Male	189	54%
		Female	161	46%
		Prefer not to say	14	4%
Age		18–30 years	217	64%
		31–45 years	105	30%
		46–60 years	14	6%
Household Income Range (₦)		Less than 50,000	18	5%
		50,000 - 100,000	41	12%
		100,001 - 200,000	57	16%
		200,001 - 500,000	147	42%
		More than 500,000	87	25%
Appliance Inventory by Type of Dwelling	**Dwelling Type**	**1-2 Persons**	**3-4 Persons**	**5 or More Persons**
	One Room/Self Contain	12%	7%	4%
	Miniflat	10%	7%	6%
	Duplex	3%	2%	3%
	3 Bedroom Flat	10%	12%	8%
	2 Bedroom Flat	5%	8%	3%

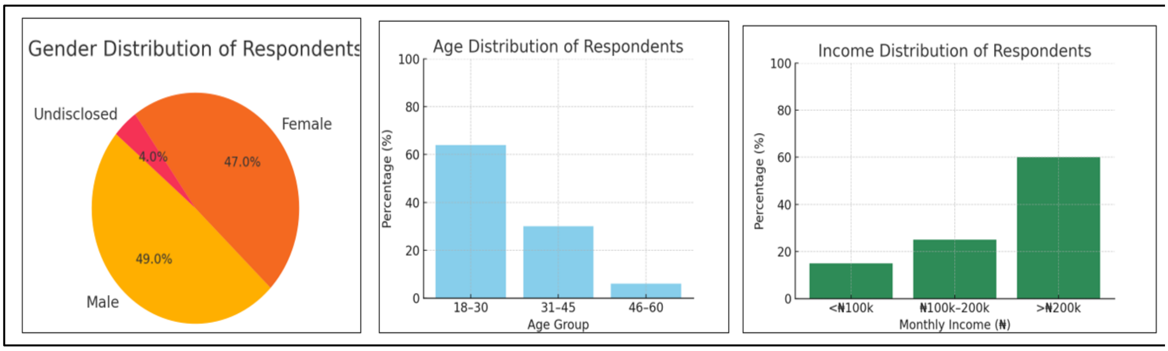

Figure 2: Demography of respondents

majority (42%) fall earn within ₦200,001 and ₦500,000 monthly, indicating a relatively high concentration in the middle-income range. This is followed by 87 respondents (25%) who earn more than ₦500,000 monthly. Conversely, the lowest income earning less than ₦50,000 has the fewest respondents (5%). These findings suggest that most households surveyed have a moderate to high income level, which may influence their consumption patterns, energy use, and access to appliances. It also suggests a possible correlation between income levels and electricity expenditure, where higher-income households are likely to have higher electricity bills due to greater appliance usage and comfort needs.

Majority (75%) of the bills paid by respondents was between ₦5,000 - ₦20,000. Households with higher electricity bills were strongly correlated with larger household sizes and income levels.

Model Results

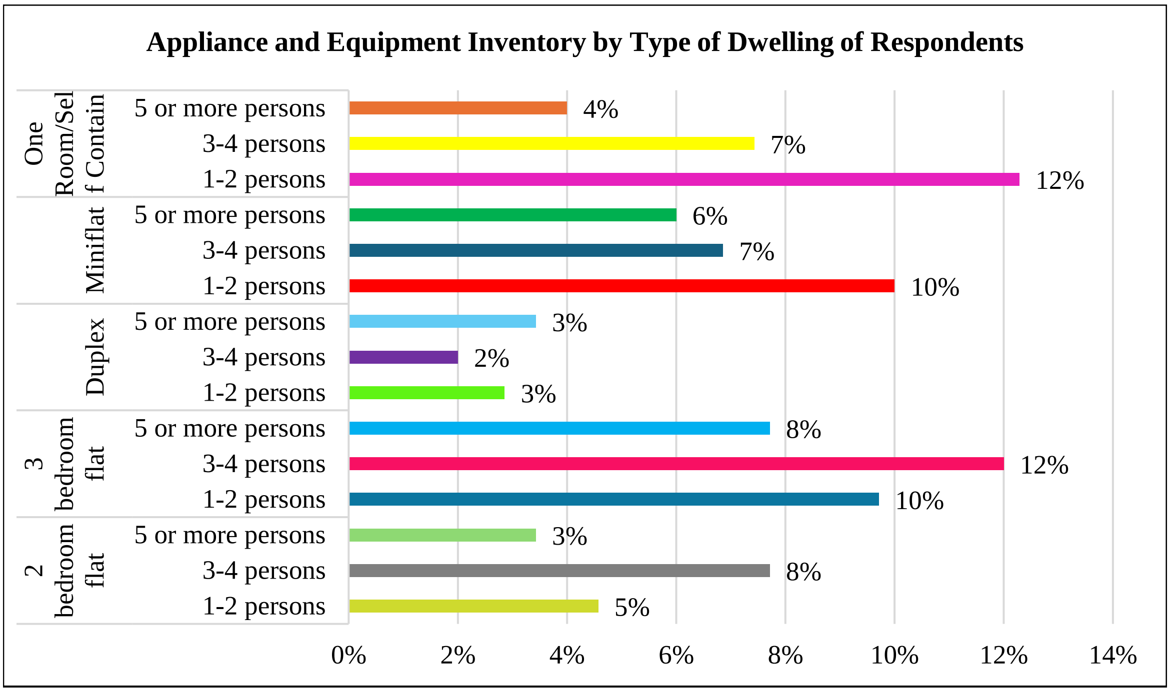

Figure 3: Appliance and equipment inventory by type of dwelling of respondents

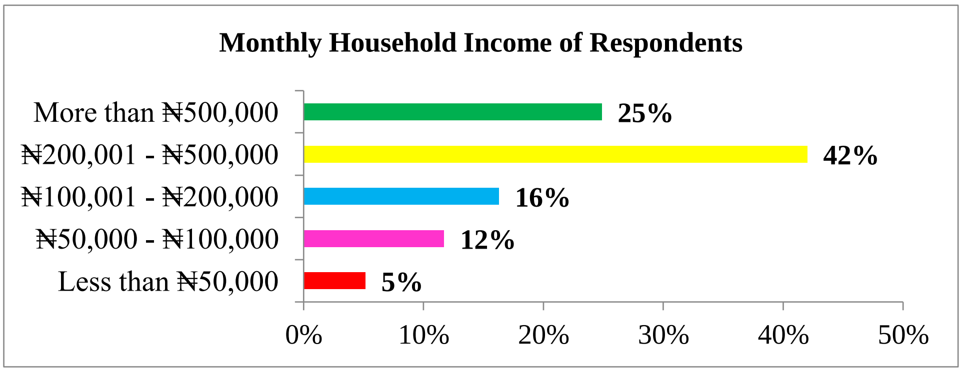

Figure 1: Monthly Household Income of Respondents

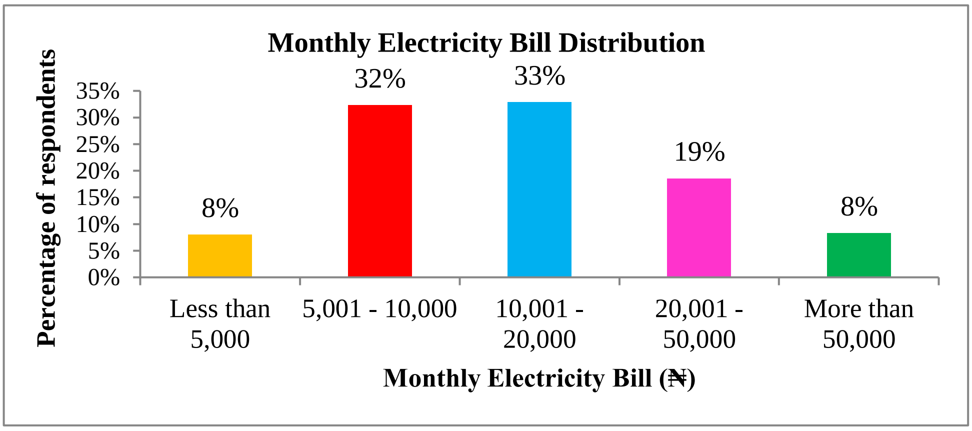

Figure 5: Monthly electricity bills

The results of the predictive models in this work are summarized using key evaluation metrics including the coefficient of determination (R^2), RMSE, and MAE. These metrics help to understand the accuracy and robustness of each model. Table 2 provides a comparative analysis of the predictive models employed in this study.

The linear regression had a R^2 value of 0.43, which indicates that it explains approximately 43% of the

Table 1: Performance of predictive models

Model	R^2 Score	RMSE	MAE
Linear Regression	0.43	1840	1300
ANN (MLP)	0.79	740	520
ARIMA	—	650	470
Hybrid ARIMA–LSTM	0.85	580	400

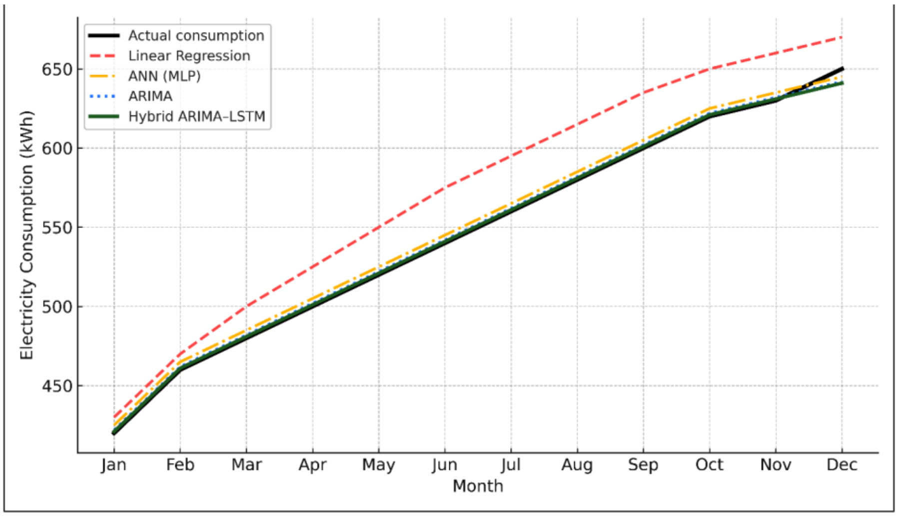

Figure 2: Comparison of actual vs predicted values for each model based on energy consumption

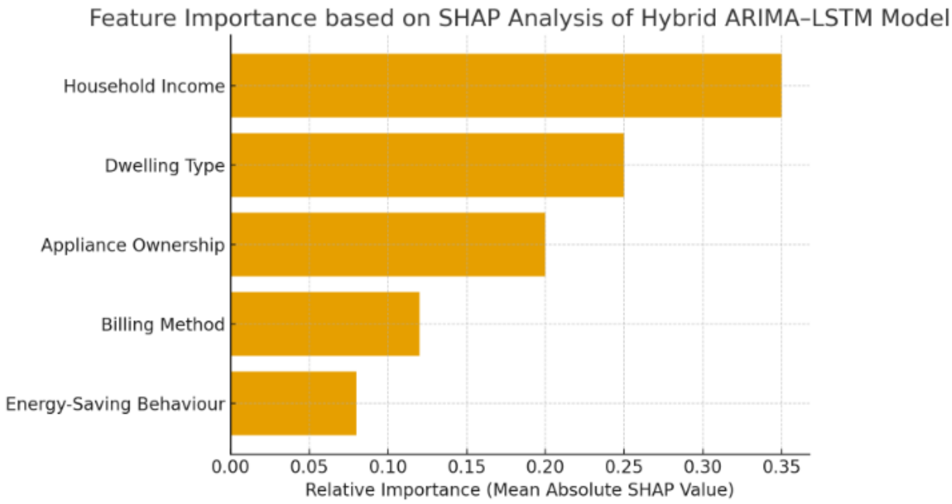

Figure 7: SHAP feature importance for the Hybrid ARIMA–LSTM model

variance in household energy consumption. This suggests a relatively weak fit to the data, likely due to its inability to capture complex, non-linear relationships inherent in residential electricity usage. Furthermore, the model recorded a high RMSE of 1,840 and a MAE of 1,300, reflecting substantial prediction errors and limited practical utility in this context. The ANN model demonstrated significantly better performance. With an R^2 value of 0.79, the ANN captured 79% of the variability in the dataset, highlighting its strength in modeling non-linear patterns and interactions among variables such as income, appliance usage, and dwelling type. The RMSE and MAE values for the ANN model were 740 and 520 respectively considerably lower than

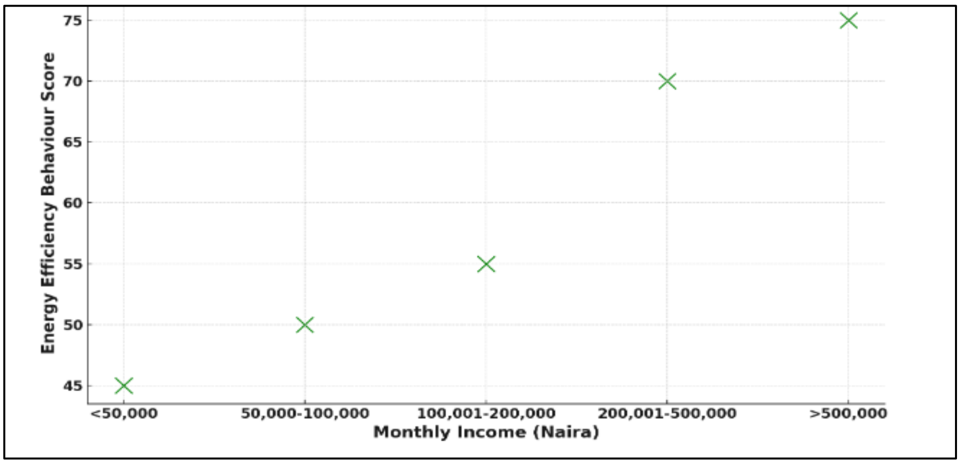

Figure 8: Scatter plot of income vs energy efficiency behaviour score

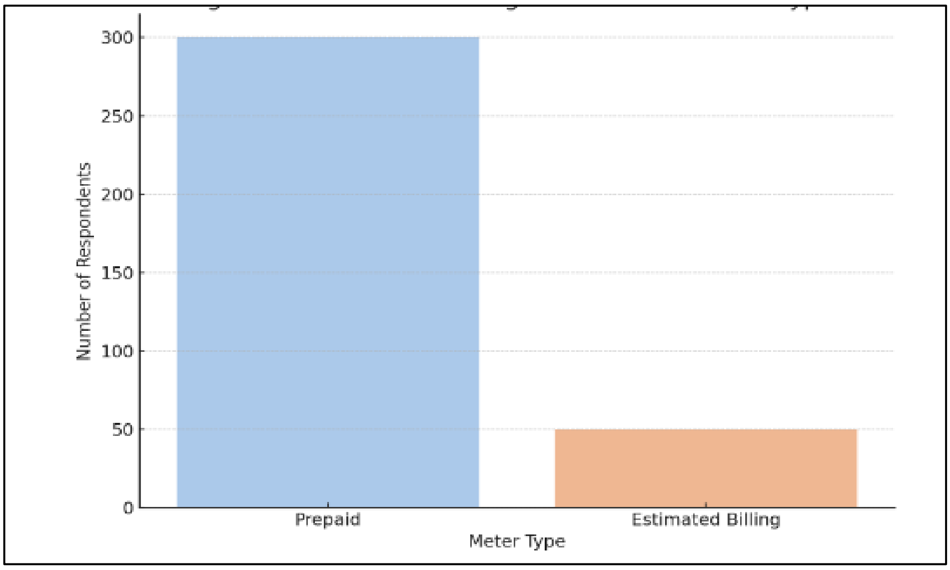

Figure 9: Bar Chart Showing Distribution of Meter Types

those of the linear model indicating improved accuracy and consistency in its predictions.

The ARIMA model also performed strongly in this study. Although it does not provide an R^2 value due to its univariate nature, the model achieved the lowest RMSE and MAE scores among the them (650 and 470 respectively). This suggests that ARIMA was particularly effective in modeling temporal trends and seasonal fluctuations in electricity consumption. However, its reliance on past consumption values and exclusion of external predictors may limit its utility in capturing broader socio-demographic influences.

Overall, the results indicate that while ARIMA is best suited for short-term, trend-based forecasting, the ANN model offers a more holistic and accurate approach for capturing complex interactions in urban household energy use. The Linear Regression model, despite its interpretability, falls short in predictive accuracy and is less suitable for the non-linear structure of the dataset. Among the models, ARIMA recorded the lowest RMSE and MAE values, indicating superior performance in short-term forecasting. The ANN model achieved the highest R^2 value, reflecting its strength in capturing complex, non-linear patterns in the data. In contrast, the linear regression model, while easy to interpret, exhibited relatively poor performance across all metrics. These findings highlight the trade-offs between model accuracy, complexity, and interpretability, which are crucial for selecting appropriate tools for energy demand forecasting.

Figure 6 compares the actual household energy consumption values with predictions generated by the forecasting models. The ARIMA model closely follows the actual consumption trends, particularly during months of peak usage. The ANN model also shows strong alignment, capturing non-linear fluctuations with minimal deviation. In contrast, the linear regression model underperforms, with wider gaps between predicted and actual values, especially during high-consumption periods. This visual comparison reinforces the superior predictive accuracy of ARIMA and ANN for short-term residential energy forecasting. Respondents in higher income brackets generally demonstrate better energy-saving practices, suggesting that income may influence access to energy-efficient technologies and awareness.

To improve the interpretability of the predictive framework, the SHAP analysis was conducted on the best-performing Hybrid ARIMA–LSTM model. The results (Figure 7) show that household income, dwelling type and appliance ownership were the most influential variables contributing to higher electricity consumption, while billing method and energy-saving behaviour had comparatively smaller effects. These outcomes are consistent with the socio-economic patterns observed in the descriptive analysis and reinforce the model's ability to capture real-world consumption behaviour. The SHAP interpretation provides a clear explanation of how individual features influenced the model's predictions, enhancing transparency and confidence in the outputs of the hybrid model.

Analysis of Energy-Saving Behaviour

Respondents showed good awareness of basic conservation strategies such as switching off appliances when not in use and using LED lighting. However, the adoption of energy-efficient appliances, such as inverter-based air conditioners and refrigerators was limited. This may be attributed to the associated costs of acquisition and installation. A strong correlation was found between income and adoption of energy-efficient behaviours (Pearson $r = 0.61$, $p < 0.05$). While income shows a positive correlation with energy-saving behaviour (Pearson $r = 0.61$, $p < 0.05$), other latent factors such as education level or environmental awareness may also play mediating roles and warrant further investigation. Figure 8 illustrates a positive relationship between household income levels and energy efficiency behaviour scores. Respondents in higher income brackets generally demonstrate better energy-saving practices, suggesting that income may influence access to energy-efficient technologies and awareness.

For the type of meter used, prepaid meters were overwhelmingly preferred (about 85%). **Figure** shows that a smaller portion of the respondents still rely on estimated billing. The prevalence of prepaid meters indicates a preference for consumption control and transparency in billing.

Conclusions

This research explored household electricity consumption patterns in Lagos, Nigeria by focusing on how socio-economic variables such as income, appliance ownership, and type of dwelling affect the use of energy. The comparative analysis of four predictive models showed that income and type of appliances owned are significant drivers of electricity consumption. Hence, in order to promote energy efficiency, there is need for behavioural interventions and demand-side management. Measures such as awareness campaigns, incentives for energy-efficient appliances, and wider adoption of prepaid metering systems can be embarked on to enhance transparency, accountability, and sustainability within the energy sector in Lagos State. Tailored education for lower-income groups and innovative financial mechanisms such as microloans, pay-as-you-save programs, and green credit facilities could facilitate broader access to clean and efficient energy technologies. Consumers can also be empowered using smart metering technologies and explainable AI like SHAP to increase transparency, trust, and regulatory acceptance. While the reliance on self-reported data and monthly time series records presented certain limitations, future research can leverage smart meter data, explore hybrid and attention-based models for greater accuracy, and expand coverage to semi-urban and rural regions. These strategies would contribute to more equitable, transparent, and resilient residential energy systems in Lagos and beyond.

Data Availability

The anonymized dataset supporting the findings of this study is available from the corresponding author upon reasonable request.

References

Adamu, M. B., Adamu, H., Ade, S. M. and Akeh, G. I. (2020). Household Energy Consumption in Nigeria: A Review on the Applicability of the Energy Ladder Model. *Journal of Applied Sciences and Environmental Management*, *24*(2), 237–244.

Bamidele, B., Omowumi, O. and Nathaniel, O. (2020). Investigating Electricity Consumption in Ogun State, Nigeria. *Journal of Engineering Studies and Research*, *26*(1), 1- 10.

Baur, L., Ditschuneit, K., Schambach, M., Kaymakci, C., Wollmann, T. and Sauer, A. (2024). Explainability and Interpretability in Electric Load Forecasting Using Machine Learning Techniques - A Review. *Energy and AI*, *16*, 1 -13.

Giacomazzi, E., Haag, F. and Hopf, K. (2023). Short-Term Electricity Load Forecasting Using the Temporal Fusion Transformer: Effect of Grid Hierarchies and Data Sources. *Proceedings of the 14th ACM International*

Conference on Future Energy Systems, 353–360.

Jang, J., Kim, B. and Kim, I. (2024). Comparative Analysis of Deep Learning Techniques for Load Forecasting in Power Systems Using Single-Layer and Hybrid Models. *International Transactions on Electrical Energy Systems*, 1, 1-22.

Ji, X., Huang, H., Chen, D., Yin, K., Zuo, Y., Chen, Z. and Bai, R. (2022). A Hybrid Residential Short-Term Load Forecasting Method Using Attention Mechanism and Deep Learning. *Buildings*, *13*(1), 72, 1 -20.

Lim, B., Arık, S. Ö., Loeff, N. and Pfister, T. (2021). Temporal Fusion Transformers for interpretable multi-horizon time series forecasting. *International Journal of Forecasting*, *37*(4), 1748–1764.

Lundberg, S. and Lee, S.-I. (2017). *A Unified Approach to Interpreting Model Predictions* (No. arXiv:1705.07874), 4768 - 4777

Nti, I. K., Teimeh, M., Nyarko-Boateng, O. and Adekoya, A. F. (2020). Electricity load forecasting: A systematic review. *Journal of Electrical Systems and Information Technology*, *7*(1), 13, 1-19.

Onatunji, O. G. (2025). Electricity consumption and industrial output: fresh evidence from economic community of West African states (ECOWAS). *Journal of Economic and Administrative Sciences,* 41(1), 381–398.

Onyenandu, N. N., Ishola, T. O. and Udu, A. A. (2025). Innovations in Electricity Infrastructure and Small Enterprises Sustainability in Lagos State, Nigeria. *NIU Journal of Social Sciences*, *11*(1), 191–205.

Sharma, G., Chandra, S., Yadav, A. K., Singh, R. and Gupta, R. (2024). A Reliable Estimation of Solar Energy Prediction Through the use of Hybrid CNN-LSTM Algorithm. *2024 2nd International Conference on Advancements and Key Challenges in Green Energy and Computing (AKGEC)*, 1–6.

Somoye, O. A. (2023). Energy crisis and renewable energy potentials in Nigeria: A review. *Renewable and Sustainable Energy Reviews*, *188*, 113794.

Suhartono, Puspitasari, I., Akbar, M. S. and Lee, M. H. (2012). Two-level seasonal model based on hybrid ARIMA-ANFIS for forecasting short-term electricity load in Indonesia: 2012 International Conference on Statistics in Science, Business and Engineering, ICSSBE 2012. *ICSSBE 2012 - Proceedings, 2012 International Conference on Statistics in Science, Business and Engineering*, 634–638.

Trading Economics. (2025). *Crude Oil Production—Countries—List*. https://tradingeconomics.com/country-list/crude-oil-production

Ubani, O., Sam-Amobi, C., Mba, E., Idu, E., Ezeama, E. and Oforji, P. I. (2024). Household electricity consumption determinants in major Nigeria cities. *Journal of Infrastructure, Policy and Development*, *8*(3), 1- 19.

Ugbehe, P. O., Diemuodeke, O. E. and Aikhuele, D. O. (2025). Electricity demand forecasting methodologies and applications: A review. *Sustainable Energy Research*, *12*(1), 1 - 32.