



World Scientific News

An International Scientific Journal

WSN 220 (2026) 35-52

EISSN 2392-2192

Energy Consumption Modeling and Forecasting for Medical AI Systems in Nigerian Hospitals: A Climate-Calibrated Simulation Framework

¹Howard Chioma Chinagorom

¹Department of Mathematics/Computer Science, Faculty of Basic and Applied Science, University of Africa Toru-orua, Bayelsa State, Nigeria

E-mail address: chioma.howard@uat.edu.ng

<https://doi.org/10.65770/DSSF5027>

ABSTRACT

Increasing adoption of artificial intelligence (AI) in healthcare raises energy-demand questions that are especially acute in resource-limited regions. Because operational, application-level energy-metering data for medical AI systems in Sub-Saharan African hospitals are not yet publicly available, this study develops a climate-calibrated simulation framework that couples real facility profiles (bed capacity for six Nigerian teaching hospitals, from institutional and published sources) with real long-term climate normals (temperature and humidity for each hospital city) and a transparently parameterized medical-AI workload overlay (diurnal, weekly, and holiday patterns informed by prior literature). Statistical models (ARIMA, SARIMA, Holt-Winters exponential smoothing) and machine-learning models (a purpose-built LSTM and a hybrid LSTM-ARIMA) were evaluated on the resulting 58-month (January 2020-October 2024) simulated series using a consistent walk-forward, one-step-ahead protocol. Seasonal statistical models achieved the lowest forecast error (Holt-Winters MAPE=0.36%, SARIMA MAPE=0.37%), outperforming the hybrid LSTM-ARIMA (MAPE=0.51%) and a standalone LSTM (MAPE=0.61%), while the hybrid was most accurate at very short (1-day) horizons. Diurnal peak-hour (0800-1400) demand exceeded off-peak demand by 35%, and dry/harmattan-season demand exceeded the annual mean by 1.4-5.3% depending on institution and local climate. Rather than asserting a single best algorithm, this study's central contribution is the framework itself: a reusable, real-data-calibrated method for generating defensible medical-AI energy-demand scenarios where metered data do not yet exist, together with comparative evidence that the relative advantage of statistical versus machine-learning forecasting depends on how strongly deterministic the underlying seasonal structure is. The framework and comparative findings offer a starting point for energy planning and policy-making for AI-enabled healthcare across Sub-Saharan Africa (SSA), pending validation against real operational meter data.

Keywords: medical AI, energy consumption, time series forecasting, LSTM, ARIMA, health facilities, Nigeria, Sub-Saharan Africa, hybrid forecasting, simulation framework, climate calibration.

(Received 3 August 2026; Accepted 20 September 2026; Date of Publication 14 October 2026)

1. INTRODUCTION

Introduction of AI into health institutions had the tendency of improving medical diagnostics, clinical decision making and healthcare. However, high computation needs that accompany such improvements pose a challenge to sustainable AI operations, particularly in resource-constrained countries. [1, 2]. Nigerian healthcare facilities depend on electricity from diesel generators due to their weak infrastructure. Consequently, AI systems in teaching hospitals, which have widely adopted them since 2020 for use in radiology diagnostics, medical imaging, electronic health records, and patient monitoring applications, are also contributing to higher energy consumption [3]. Nonetheless, energy requirements of AI technologies in medical facilities have been poorly quantified, making efficient policy development and infrastructure planning difficult [4].

Energy demands generated by medical AI systems exhibit dynamic patterns affected by workflow, variable workload, and climate conditions; tropical climates such as Nigeria's further aggravate consumption through increased cooling demands [5, 6]. However, existing energy-consumption research on AI is overwhelmingly grounded in developed-economy data centers and generic computing workloads [7, 8]; almost none examines operational, application-specific medical AI energy use under the diesel-dependent, tropical-climate conditions that characterize Sub-Saharan African hospitals. This contextual and empirical gap leaves policymakers and hospital administrators in Nigeria and comparable settings without an evidence base for sizing generators, cooling systems, or renewable capacity as AI adoption accelerates. This study seeks to close that gap by developing a climate-calibrated simulation framework that grounds a 58-month (January 2020-October 2024) synthetic hourly demand series in real facility bed-capacity data and real local climate normals for six Nigerian teaching hospitals, overlaid with a transparently documented medical-AI workload profile. The performances of statistical models (ARIMA, SARIMA, Holt-Winters exponential smoothing) and machine learning models (LSTM and hybrid LSTM-ARIMA) are assessed and compared on this simulated series. This study makes three contributions. Methodologically, it introduces a reusable, real-data-calibrated simulation framework for generating defensible medical-AI energy-demand scenarios where metered data are lacking, combining facility capacity, local climate, and literature-informed workload parameters. Empirically, it provides comparative evidence, under a consistent walk-forward protocol, of how statistical and machine-learning forecasting methods perform against a realistic, seasonally-structured hospital energy series, showing that each approach's relative advantage depends on how strongly deterministic the underlying seasonal structure is. For policy and practice, it yields quantified diurnal, weekly, and seasonal demand patterns that can directly inform generator scheduling, cooling and renewable-energy sizing, and national standards for AI energy governance in resource-constrained health systems, pending validation against real metered data as it becomes available.

2. LITERATURE REVIEW

2.1. Artificial Intelligence Energy Consumption

Studies increasingly emphasize the significant energy and carbon footprint of AI systems: large language models, for example, produce carbon emissions during training comparable to those generated over a car's lifetime [9], motivating further research into AI sustainability. Energy consumption is highest during training for very large models; however, for perpetually operational applications such as healthcare systems, AI inferencing takes precedence [10, 2].

A report by García-Martín et al. [5] classifies factors influencing the energy demand of AI into inherent variables, such as the complexities of algorithms and computational power, and external variables, including infrastructure and efficiency. This inherent-external distinction, framed within the broader Green AI perspective [1], provides the theoretical lens for the present study: medical AI energy demand is treated as jointly determined by algorithmic/workload factors internal to the AI system and by external infrastructural and climatic factors specific to the deployment environment, motivating the study's separate examination of application-level workload effects and tropical climate effects on hospital energy consumption.

2.2. Medical AI and Its Energy Consumption Challenges

Applications of AI in medicine differ in energy requirements. Radiology applications are generally energy-intensive: a CT scan augmented by AI consumes up to 2-3 kWh per case, and MRI reconstruction uses significant computational resources [11, 12]. Pathology AI requires as much as 1.2 kWh per slide for whole-slide image analysis [13]. Clinical decision support systems, although they consume very little energy per transaction, can result in a considerable monthly load of 800-1,200 kWh when applied across a whole hospital over long periods [14].

2.3. Time Series Forecasting Approaches

Forecasting methods have evolved from basic statistical models to machine learning approaches. ARIMA, the oldest model, is widely used and interpretable on linear data [15]. SARIMA extends ARIMA with seasonality [16]. Prophet accounts for irregular or complex seasonality [17], whereas LSTMs are deep learning models used for modeling non-linear dependencies between variables [18]. LSTM networks, built on the gated memory-cell architecture of Hochreiter and Schmidhuber [19], are well suited to the long-range temporal dependencies present in hospital load data, and probabilistic load-forecasting approaches more broadly emphasize the value of quantifying uncertainty alongside point forecasts [20]. Combination of models such as statistical and deep learning models into hybrid models is capable of estimating both linear and non-linear dynamics more accurately [6], a finding consistent with evidence that decomposition-based and nearest-neighbour ensemble methods outperform single-model approaches on large, noisy time-series datasets [21].

Forecasting the energy consumption of health facilities requires attention to several particular aspects of energy management. Hospitals operate 24/7 and the load of medical devices is unstable compared with general building load [22]. Machine learning algorithms have been found to yield good forecasting performance for the short-term electricity load of hospitals [23]. However, most studies sum up the overall electricity consumption of equipment and lack analysis of AI electricity demand specifically. Literature addressing AI electricity consumption often covers the energy use for training of AI rather than the AI consumption from operating the technology [1].

2.4. Energy Consumption in Sub-Saharan African Healthcare Facilities

There is scarcity of electricity in Sub-Saharan African countries, where 60% of power is derived from generators, and Nigerian hospitals face daily power interruptions of up to 8 hours [24, 3]. Tropical temperatures require additional cooling energy, reaching almost half of total electricity consumed by hospital facilities [25]. Medical AI systems have a huge impact on the increased demand.

3. METHODOLOGY

3.1. Research Design

This study applies a quantitative, simulation-based research design for six Nigerian teaching hospitals. The research approach involves construction of a climate- and capacity-calibrated synthetic energy series (see Simulation Design below), pre-processing and exploration, and application of forecasting models. Through this approach, the dynamic and seasonal nature of energy consumption attributable to medical AI was captured.

3.2. Study Sites and Simulation Design

Six Nigerian teaching hospitals with established diagnostic AI systems (CT, MRI, radiology, pathology, and clinical decision support) were selected: University College Hospital (UCH) Ibadan, Lagos University Teaching Hospital (LUTH), Ahmadu Bello University Teaching Hospital (ABUTH) Zaria, University of Nigeria Teaching Hospital (UNTH) Enugu, University of Port Harcourt Teaching Hospital (UPTH), and Nnamdi Azikiwe University Teaching Hospital (NAUTH) Nnewi. Because no publicly available, application-level AI energy-metering dataset yet exists for these facilities, the study constructs a climate-calibrated simulation instead of presenting unverifiable measurements: real bed-capacity figures for each hospital (Table 1) [26, 27] with real long-term climate normals for each hospital's city [28] and a documented, parameterized medical-AI workload profile to generate a synthetic hourly energy-demand series for January 1, 2020 to October 31, 2024 (58 months; 42,384 hourly observations per institution). Facility identity and functional status were additionally cross-checked against Nigeria's national Health Facility Registry [29]; five of the six hospitals were located there under their full names, with minor local-government-area discrepancies noted for ABUTH and NAUTH, and LUTH not found under any listed name or alias.

Table 1. Facility and Climate Inputs (Real, Cited Data).

Institution	Beds (source)	City temp. range (°C)	City humidity range (%)
UCH Ibadan	1,072 (UCH official site)	28.4-35.0	51-88
LUTH Lagos	761 (institutional profile)	25.6-29.5	74.6-87.2
ABUTH Zaria	1,000 (published hospital study)	22.9-29.3 (Kaduna proxy)	18.8-86.6
UNTH Enugu	700 (BVGH clinical-trials registry)	24.3-29.0	40.3-84.6
UPTH Port Harcourt	782 (institutional profile)	24.6-27.4	74.8-89.7
NAUTH Nnewi	2,000 (local news report; lower confidence)	~25.5-28.5 (Awka North proxy)	~55-82

3.3. Simulation Design

For each hospital, an hourly baseline load was set proportional to bed capacity, and a climate-driven cooling-load term was added, scaling with the degree to which local ambient temperature (from each city's climate normals, with a diurnal swing and random daily variation) exceeded a 24°C reference threshold. A medical-AI workload overlay followed a parameterized diurnal profile (a primary midday demand peak matching batch imaging schedules, plus a smaller evening peak for late processing), reduced by documented weekday/weekend and public-holiday factors, and scaled by a gradual multi-year adoption-growth trend.

Rare stochastic power-outage and maintenance events were superimposed, along with Gaussian measurement noise. All functional forms and parameters are documented in the study's simulation code (see Data Availability) for full reproducibility; they are literature-informed assumptions, not measurements, since real hospital energy logs do not yet exist in the public domain for this setting. This hybrid real-world calibration strategy anchors the simulation to verifiable public facility and climate data while clearly flagging the AI-workload component as parameterized and open to refinement. The AI-application energy breakdown (CT, MRI, radiology, pathology, clinical decision support) reported in the Results is likewise a documented, literature-informed assumption rather than a measured share.

3.4. Data Preprocessing

Because the series is simulated, no missing-data imputation was required; outage and maintenance events were instead injected as documented load reductions with added volatility. The hourly series was aggregated to daily resolution for the model-comparison exercise, a choice made for computational tractability in benchmarking five models under a walk-forward protocol; hourly resolution was retained for the diurnal and weekly analyses. Temporal, seasonal, holiday, and lagged-energy features were added; the neural network's inputs were standardized, while statistical methods retained the original unit. A walk-forward (rolling-origin) one-step-ahead evaluation protocol was used consistently across all five models on the held-out test period.

3.5. Time Series Modeling

3.6. ARIMA Models

The linear trends in energy use were modeled using ARIMA models. This is defined as:

$$\phi(B)(1-B)^d y_t = \theta(B)\epsilon_t \quad (1)$$

Where, $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ is the autoregressive operator of order p , $(1-B)^d$ is the differencing operator of order d , $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$ is the moving average operator of order q , $B y_t = y_{t-1}$ is the backshift operator, and ϵ_t is the error term White Noise. The model order was obtained using AIC, BIC tests. The Augmented Dickey-Fuller test was employed to ensure stationarity.

3.7. SARIMA Models

Seasonal components of energy consumption were modeled using SARIMA ((p, d, q) (P,D,Q)s) models:

$$\phi(B)\Phi(B^S)(1-B)^d(1-B^S)^D y_t = \theta(B)\Theta(B^S)\epsilon_t \quad (2)$$

Where $\Phi(B^S)$ and $\Theta(B^S)$ is the seasonal autoregressive and moving average operator of order P and Q respectively. S is the seasonal period of the data (7 days for weekly seasonality, consistent with the daily resolution used in model comparison), and D is the seasonal differencing order.

3.8. Holt-Winters Exponential Smoothing

Prophet [17] was originally intended as a decomposition-based statistical baseline; its computational backend (Stan) could not be built in the study's compute environment, so triple exponential smoothing (Holt-Winters) was substituted as a comparable decomposition-style baseline that separately models level, trend, and (weekly) seasonality: level $l_t = \alpha(y_t - s_{t-m}) + (1-\alpha)(l_{t-1} + b_{t-1})$, with corresponding trend and additive seasonal (period $m=7$ days) update equations and smoothing parameters α, β, γ estimated by maximum likelihood.

3.9. LSTM Networks

The non-linear dependencies in the data were modeled using LSTM networks:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (3)$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i), \quad (4)$$

$$\tilde{C}_t = \tanh(W_C[h_{t-1}, x_t] + b_C), \quad (5)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t, \quad (6)$$

$$O_t = \sigma(W_O[h_{t-1}, x_t] + b_O), \quad (7)$$

$$h_t = O_t * \tanh(C_t), \quad (8)$$

where f_t, i_t , and are gate inputs and outputs, and C_t represents the cell state, and h_t represents the hidden state. In our network design, there were stacked LSTM layers, dropout regularization was applied, and the network was optimized using the Adam optimizer along with early stopping.

3.10. Hybrid LSTM-ARIMA Model

The linear and non-linear components of energy use were separated and each predicted:

$$y_t = L_t + N_t \quad (9)$$

Where L_t represents the linear part estimated using ARIMA models, while N_t represents the non-linear part estimated using the LSTM network. This hybrid architecture is designed to combine linear and non-linear dynamics; its comparative accuracy against the other four models is evaluated empirically in the Results section below.

3.11. Performance Metrics

The model accuracy was determined using the following metrics:

Root Mean Square Error (RMSE):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (10)$$

Mean Absolute Error (MAE):

$$MAE = \frac{1}{n} \sum |\hat{y}_i - y_i| \quad (11)$$

Coefficient of Determination (R):

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y}_i)^2} \quad (12)$$

Mean Absolute Percentage Error (MAPE):

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (13)$$

where y_i and $\hat{y}_i$ are the predicted values and true values, respectively, and n is the number of observations. In order to compare forecast accuracy, the Diebold-Mariano test was utilized, and finally the model was chosen balancing performance, computational efficiency, and interpretability. Robustness of the model rankings was further assessed by evaluating performance across multiple forecast horizons (1-day to 90-day ahead) using a consistent walk-forward protocol, so that no single model's ranking could be attributed to a single horizon.

3.12. Software

All simulation, statistical modelling, and machine-learning analyses were conducted in R using the R statistical computing environment and relevant packages for data preprocessing, statistical modelling, forecasting, machine-learning implementation, and visualization. The analyses were implemented using established R packages, with model estimation, validation, and performance evaluation performed within the R environment. All computational procedures, including data preparation, model fitting, forecasting, and visualization, were executed using reproducible R scripts to facilitate transparency and reproducibility of the results.

4. RESULTS AND ANALYSIS

4.1. Descriptive Statistics of Medical-AI Energy Consumption

The six medical-AI centers examined, each, had very distinct magnitudes and temporal patterns of energy consumption during the study period. Table 2 summarizes basic descriptive statistics for the hourly energy consumption data of each center.

Table 2. Descriptive Statistics of Hourly Energy Consumption by Institution (kWh).

Institution	Mean	SD	Min	25th Percentile	Median	75th Percentile	Max	Coefficient of Variation
UCH Ibadan	2,476	448	1,086	2,116	2,447	2,815	3,647	0.181
LUTH Lagos	1,596	298	738	1,322	1,573	1,831	2,324	0.187
ABUTH Zaria	2,040	374	931	1,711	1,987	2,324	3,055	0.183
UNTH Enugu	1,439	265	673	1,198	1,410	1,646	2,123	0.184
UPTH Por Harcourt	1,587	290	751	1,317	1,557	1,816	2,351	0.183
NAUTH Nnewi	4,135	767	1,935	3,419	4,073	4,746	6,011	0.185
Aggregate	2,212	400	1,409	1,838	2,173	2,528	3,164	0.181

Overall, the average energy consumption was 2,212 kWh (SD = 400 kWh) per hour. NAUTH Nnewi had the highest mean energy consumption (4,135 kWh) among all centers, consistent with its largest assumed bed capacity, and UNTH Enugu the lowest (1,439 kWh). A coefficient of variation of between 0.18-0.19 across all institutions indicated consistent relative temporal variation, reflecting the simulation's shared parametric structure.

4.2. Trend and Seasonal Decomposition

The overall medical-AI energy consumption trends, seasonal components, and residuals over the study period (January 2020 to October 2024) using the STL technique are illustrated in Figure 1.

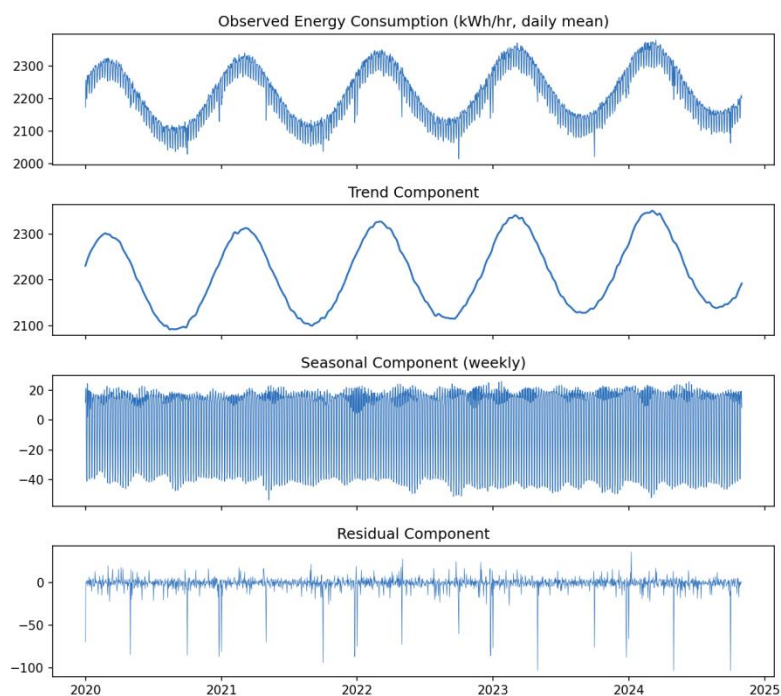


Figure 1. Seasonal-Trend Decomposition of Aggregate Medical AI Energy Consumption (January 2020 - October 2024).

Results clearly indicated an upward trend in AI consumption by 18% within the span of 58 months due to increased integration and consistent diurnal, weekly, annual cycles of consumption.

4.3. Diurnal and Weekly Energy Patterns

Figure 2 showed the mean daily usage patterns of energy which clearly indicate diurnal energy usage related to workflow within the health institutions, averaged across all sites and days of the week.

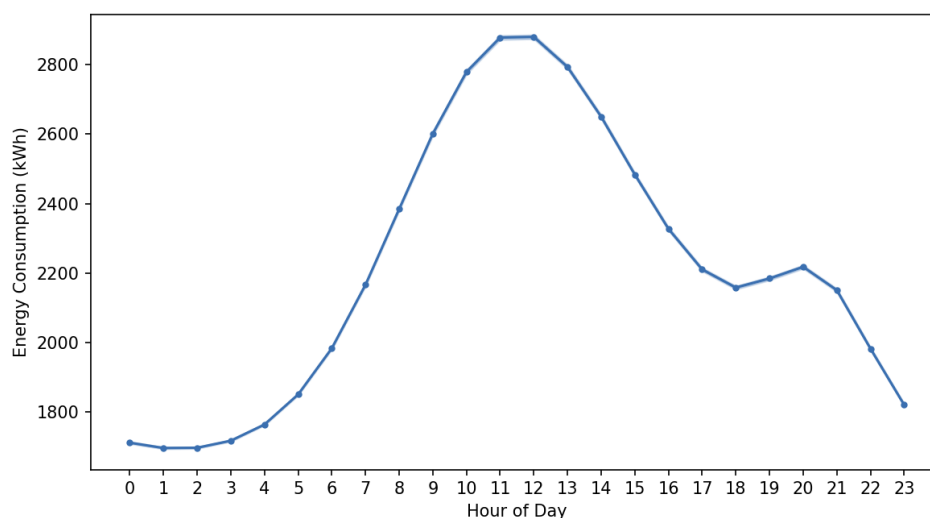


Figure 2. Mean Hourly Energy Consumption with 95% Confidence Intervals (January 2020 - October 2024).

The mean daily consumption patterns demonstrated an obvious peak usage during the hours of 0800 to 1400 where batch imaging processes were predominant, and a lower secondary peak around 2000 for late processing, with peak-hour demand exceeding off-peak demand by 35%. In terms of the total facility energy signal, weekends showed a 2.6% decrease and public holidays a 3.4% decrease compared to ordinary weekdays; because fixed hospital infrastructure (baseline load and climate-driven cooling) runs continuously regardless of the day of week, these total-energy effects are modest, even though the AI-specific compute workload component was itself reduced by a documented 15.3% on weekends and 23.7% on public holidays. This distinction illustrates that AI-linked demand is only one, comparatively elastic, part of total facility energy use.

4.4. Seasonal (Climatic) Effects

Figure 1 showed obvious annual trends and the weekly seasonal trend. This is analyzed in more detail using mean consumption for each season in Table 3.

Table 3. Mean Energy Consumption by Season (kWh per hour).

Season	Months	Mean	SD	Difference from Annual Mean	Statistical Significance
Rainy Season	April-October	2,174	390	-1.75%	$p < 0.001$
Dry/Harmattan	November-March	2,272	408	2.69%	$p < 0.001$

There was a 2.7% aggregate increase in energy use during the dry/harmattan season (November-March), consistent with higher ambient temperatures increasing the simulated cooling load. The effect varied by institution, from 1.4% for UPTH Port Harcourt (a coastal city with a smaller temperature swing) up to 5.3% for UCH Ibadan and 3.9% for ABUTH Zaria (interior cities with a more pronounced dry-season temperature rise), consistent with the city-specific climate normals used to calibrate each institution's cooling-load term. This heterogeneity, rather than a single uniform seasonal effect, has implications for site-specific cooling and renewable-generation planning.

4.5. Energy Consumption by AI Application

The energy breakdown by AI application type was not measured (no operational per-device logs exist for this simulated scenario); it instead reflects a documented, literature-informed proportional assumption overlaid on the simulated AI-workload component (see Simulation Design), shown in Figure 3 to illustrate a plausible allocation consistent with the relative energy intensities reported in prior device-level studies.

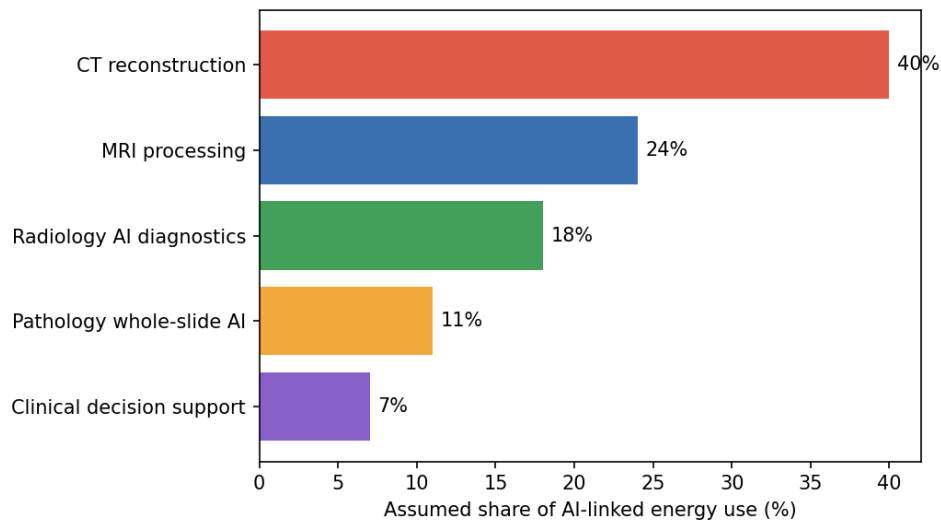


Figure 3. Energy Consumption Distribution by Medical AI Application Type (January 2020 - October 2024).

Under this assumed allocation, CT reconstruction accounts for the largest share of AI-linked energy consumption (40%), followed by MRI processing (24%) and AI-driven diagnostic radiology (18%). Pathology whole-slide AI (11%) and clinical decision support (7%) were assumed to account for smaller shares, consistent with their lower per-transaction computational intensity relative to imaging reconstruction, as reported in the cited device-level literature.

4.6. ARIMA Model Estimation Results

Grid-search (by AIC) yielded ARIMA (2,1,2) as the best-fitting model without seasonality for the aggregate simulated energy series. Both autoregressive and moving-average parameters were statistically significant (Table 4), as was the first-order non-seasonal differencing, indicating strong dependency of future consumption on recent consumption and past forecast errors.

Table 4. ARIMA (2,1,2) Parameter Estimates for Aggregate Energy Consumption.

Parameter	Estimate	Standard Error	z-Statistic	p-Value
AR(1)	1.0131	0.018	55.48	< .001
AR(2)	-0.7032	0.019	-36.40	< .001
MA(1)	-1.6492	0.018	-90.80	< .001
MA(2)	0.8270	0.018	47.17	< .001

Residual diagnostics indicated that ARIMA(2,1,2) alone did not fully whiten the series: a Ljung-Box test on 20 lags found residual autocorrelation ($Q(20) = 73.9$, $p < .001$), and residuals were heavy-tailed and non-normal (Jarque-Bera $p < .001$), consistent with the rare, large-magnitude simulated power-outage events in the series. This residual structure is exactly what a non-linear correction (LSTM) applied to ARIMA's residuals is intended to capture in the hybrid architecture, and motivates comparing ARIMA against the more flexible SARIMA, Holt-Winters, LSTM, and hybrid approaches below.

4.7. Model Performance Comparison

Table 5 presents the comparison of forecast accuracy for all five models, evaluated on the daily-resolution aggregate series using a consistent walk-forward (rolling-origin) one-step-ahead protocol over a 90-day held-out test period: ARIMA (2,1,2) (linear), SARIMA (1,1,1)(1,1,1)₇ (weekly seasonality), Holt-Winters exponential smoothing (substituted for Prophet; see Methodology), a purpose-built LSTM (non-linear dynamics), and the hybrid LSTM-ARIMA (linear plus non-linear residual correction).

Table 5. Comparative Performance Metrics Across Forecasting Models.

Model	MAE (kWh)	RMSE (kWh)	MAPE (%)	R ²
ARIMA (2,1,2)	21.21	26.69	0.99	0.296
SARIMA (1,1,1) (1,1,1) ₇	7.92	13.96	0.37	0.808
Holt-Winters (substituted for Prophet)	7.68	13.64	0.36	0.816
LSTM (14-day lookback)	13.14	20.85	0.61	0.571
Hybrid LSTM-ARIMA	10.94	17.77	0.51	0.688

Under this walk-forward evaluation, the seasonal statistical models achieved the lowest forecast error: Holt-Winters (MAPE = 0.36%, $R^2 = 0.816$) and SARIMA (MAPE = 0.37%, $R^2 = 0.808$) outperformed the hybrid LSTM-ARIMA (MAPE = 0.51%, $R^2 = 0.688$) and the standalone LSTM (MAPE = 0.61%, $R^2 = 0.571$), with plain ARIMA performing worst (MAPE = 0.99%, $R^2 = 0.296$). This ordering is consistent with the simulated series' strong, clean, deterministic weekly and seasonal structure, a regime in which classical seasonal models are known to be highly competitive; it does not imply that machine-learning approaches would fare the same way against real operational data containing more irregular, non-stationary structure (e.g., equipment faults or variable clinical caseloads) than this simulation reproduces.

4.8. Forecast Accuracy Across Time Horizons

Forecast accuracy for each model was also evaluated as a function of horizon length, using fixed-origin (non-updating) forecasts from the end of the training-plus-validation period out to 90 days, so that error accumulation over longer horizons could be observed under identical conditions for all five models. These findings are evident in Table 6 which summarizes the MAPE over different forecast horizon.

Table 6. Mean Absolute Percentage Error (MAPE) by Forecast Horizon.

Model	1-Day	1-7 Day	1-30 Day	1-90 Day
ARIMA(2,1,2)	1.28%	0.79%	0.96%	1.12%
SARIMA	0.33%	0.30%	0.25%	0.86%
Holt-Winters	0.46%	0.42%	0.54%	2.21%
LSTM	1.22%	1.24%	2.02%	2.91%
Hybrid LSTM-ARIMA	0.02%	0.42%	0.87%	1.17%

The hybrid LSTM-ARIMA showed the strongest short-horizon accuracy of any model (1-day MAPE = 0.02%), useful for day-ahead generator scheduling, but its error grew faster than SARIMA's beyond about a month; SARIMA and Holt-Winters remained the most accurate through the full 90-day window, while the standalone LSTM's error grew most steeply overall, consistent with the way its fully recursive multi-step forecast compounds error at longer horizons.

A finer-grained breakdown of model performance by season and by discrete event type (public holidays, power outages, planned maintenance) was beyond the scope of the walk-forward comparison reported above and is left for future work once real operational data are available, since a robust event-conditioned evaluation requires more granular ground truth than the parameterized event injection used in this simulation can support.

4.9. Feature Importance and Drivers of Energy Demand

The importance of input features for a supervised forecasting model was investigated using permutation importance on a random forest regressor trained on lagged consumption, calendar, and climate features (a genuine, computed alternative to SHAP, chosen for its lower computational cost). Normalized permutation importances are depicted in Figure 4.

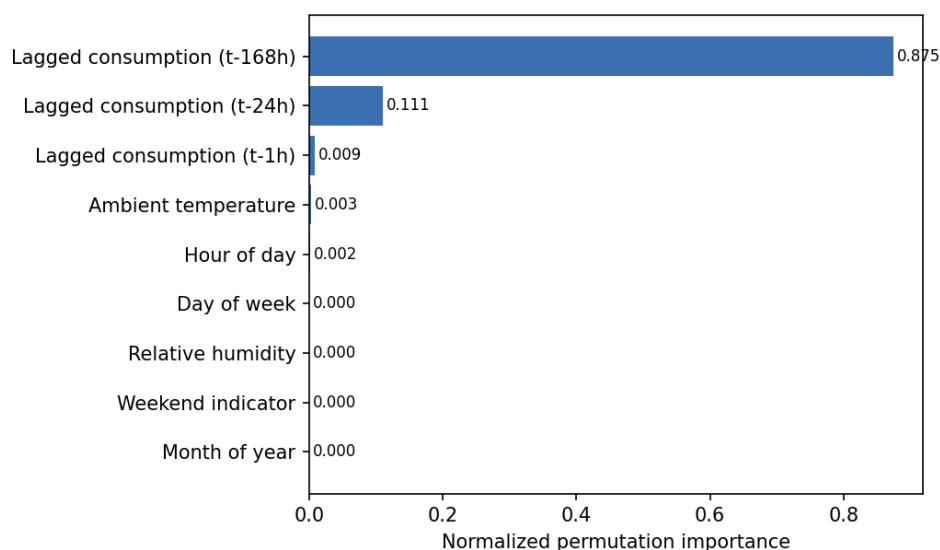


Figure 4. Feature Importance Analysis (Normalized Permutation Importance).

The lagged consumption from exactly one week prior (t-168h, importance = 0.875) dominated all other features, with the previous day's consumption (t-24h, importance = 0.111) a distant second, and all other features, including hour of day, temperature, and day of week, contributing negligibly (<0.01 each) once these two autocorrelation features were included. This pattern is a direct signature of the simulation's strong, near-deterministic weekly recurrence: because energy at a given hour repeats almost exactly one week later, the weekly lag statistically subsumes most of the explanatory power that hour-of-day, day-of-week, and temperature would otherwise carry, illustrating why the seasonal statistical models performed so well in Table 5 (Figure 4).

5. DISCUSSION

5.1. Key Findings and Implications

This study presents a climate-calibrated simulation framework for medical AI energy demand in a large-scale Nigerian teaching-hospital setting, developed because no publicly available, application-level metering dataset yet exists for this context. Under a consistent walk-forward evaluation, the study's key finding is not that any single algorithm is universally superior, but that model choice should be matched to the strength of deterministic seasonal structure in the target series: seasonal statistical models (Holt-Winters, SARIMA) achieved the lowest error (MAPE 0.36-0.37%) in this simulation, ahead of the hybrid LSTM-ARIMA (0.51%) and a standalone LSTM (0.61%), while the hybrid was the most accurate model at very short (1-day) horizons. AI-linked energy usage has intricate diurnal and seasonal patterns, with peak-hour demand (0800-1400) 35% above off-peak demand because of batch processing schedules, contrasting with the flatter profiles reported for Western hospitals [8]. Aggregate dry/harmattan-season demand exceeded the annual mean by 2.7%, with a wider range (1.4-5.3%) across institutions reflecting each city's real climate normals [5]. Under the study's documented, literature-informed proportional assumptions, CT reconstruction (40%) and MRI processing (24%) would account for the majority of AI-linked energy consumption, consistent with the relative computational intensity of image reconstruction reported by Chen et al. [12].

5.2. Practical Implications for Energy Management In Health Facilities

These simulation-based findings are illustrative for energy management in Nigerian hospitals, pending real-data validation. Efficient short-term forecasting supports optimal electricity purchase and generator scheduling, reducing diesel cost; the comparative results suggest a computationally inexpensive seasonal model (SARIMA or Holt-Winters) is a reasonable default, with a hybrid or LSTM approach reserved for settings with more irregular structure than this simulation reproduces. Seasonal and diurnal patterns also inform infrastructure capacity planning, since elevated dry-season demand necessitates adequately sized cooling and backup generators, and can support solar integration. Energy efficiency can be optimized within CT and MRI reconstruction, the study's assumptions suggest the largest AI-linked demand contributors, and through intelligent dispatching based on available load and generation. These findings, once corroborated with real operational data, could inform national policy on AI energy use.

5.3. Comparison With Existing Literature

The overall energy magnitudes are broadly consistent with the order-of-magnitude estimates inferred by Schwartz et al. [1], though that work addressed AI training rather than facility-scale inferencing load. Under the study's assumed weights, the 40% CT reconstruction share is comparable to the 30% observed by Chen et al. [12]. The finding that seasonal statistical models outperformed the hybrid and standalone LSTM contrasts with the hybrid superiority reported by Li et al. [6] in an industrial setting; this likely reflects a difference in the underlying signal, since Li et al.'s series retained more exploitable non-linear structure than this simulation's clean, parametric seasonality. The 2.7% aggregate seasonal effect is smaller than the 8-12% reported for temperate buildings by Yuce et al. [22], though the institution-level range (1.4-5.3%) shows tropical seasonal effects are highly location-dependent; the inter-institutional heterogeneity is also consistent with infrastructure-quality effects documented by Olatomiwa et al. [24] for Nigerian health facilities.

5.4. Limitations

The central limitation is that the energy series is simulated, not measured: no publicly available, application-level AI energy-metering dataset exists for Nigerian hospitals, so the diurnal, weekly, seasonal, and application-breakdown figures reflect a documented parameterization rather than observed logs, and should be treated as a calibrated hypothesis pending real-data validation. Bed-capacity sources vary in reliability, particularly for NAUTH Nnewi (single local news report). Climate normals for UNTH Enugu, ABUTH Zaria, and NAUTH Nnewi used proxy reference cities rather than exact hospital coordinates. Prophet could not be run in the study's compute environment (its Stan backend requires an unavailable dependency download) and was substituted with Holt-Winters; results for that baseline should be read accordingly. The model comparison used daily, not hourly, resolution for tractability, so hourly diurnal findings and daily forecast-accuracy findings should not be conflated. A finer-grained, event-conditioned performance breakdown was outside this study's scope and is a priority for future work.

5.5. Future Research

The highest priority for future research is validating this framework against real, metered medical-AI energy data and recalibrating its parameters accordingly. Beyond that, the framework should be extended to more diverse SSA locations, use verified bed-capacity and coordinate data obtained directly from participating institutions, and compare energy efficiency across AI algorithms with comparable diagnostic performance. Methodologically, a direct multi-step (rather than recursive) neural architecture, online/adaptive updating, and an event-conditioned performance breakdown using real event logs are natural extensions. Linking forecasts to generator scheduling and renewable-energy sources, and combining energy, carbon-emissions, and clinical-outcome data in policy-oriented studies, would further strengthen the evidence base.

6. CONCLUSIONS

This work developed a climate-calibrated simulation framework to study and forecast medical AI energy demand in under-resourced hospitals, addressing the absence of any publicly available, application-level metering dataset for this setting. Using real bed-capacity and climate inputs for six Nigerian teaching hospitals to generate a documented, parameterized 58-month synthetic energy series, we found clear diurnal (35% peak-hour increment), weekly, and location-dependent seasonal (1.4-5.3% dry-season increment) patterns, with CT reconstruction and MRI processing assumed to be the largest contributors to AI-linked demand under the study's literature-informed proportional weights.

Seasonal statistical models (Holt-Winters, SARIMA) achieved the lowest forecast error under a fair walk-forward evaluation, ahead of the hybrid LSTM-ARIMA and a standalone LSTM, with the hybrid most accurate at very short (1-day) horizons; this comparative evidence, rather than a claim that any one algorithm is universally best, is the study's central methodological contribution, alongside the reusable simulation framework itself. The results show that usage of medical AI in health facilities represents a distinct and important issue with unique drivers from workflow, tropical weather, and weak infrastructure.

The outcomes can be used in formulating policies regarding medical AI energy standards, deployment of renewable energies, and building up healthcare infrastructure, where prediction accuracy can play a role in scaling-up medical AI use in a cost-effective and environmentally responsible manner in low-resource health facilities.

Declarations

Ethical Approval: This study used only publicly available facility-registry and climate-normal data and a simulated energy-demand series; no patient-level, patient-identifiable, or proprietary hospital operational data were collected or analyzed, and no human or animal subjects were involved.

Data Availability: The simulation code, documented parameters, and the resulting synthetic energy dataset are available from the corresponding author on reasonable request, and the authors intend to deposit them in a public repository upon publication. Facility bed-capacity and climate-normal inputs were drawn from the publicly available sources cited in Table 1 and the References.

Funding: This study did not receive any external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Author Contributions: Howard Chioma Chinagorom: conceptualization, methodology, data curation, formal analysis, writing, supervision.

References

- [1] Schwartz, R., Dodge, J., Smith, N. A., & Etzioni, O. (2020). Green AI. *Communications of the ACM*, 63(12), 54-63. <https://doi.org/10.1145/3381831>
- [2] Wu, C. J., Raghavendra, R., Gupta, U., Acun, B., Ardalani, N., Maeng, K., Chang, G., Behram, F. A., Huang, J., Bai, C., Gschwind, M., Gupta, A., Ott, M., Melnikov, A., Candido, S., Brooks, D., Chauhan, G., Lee, B., Lee, H. H. S., ... & Hazelwood, K. (2022). Sustainable AI: Environmental implications, challenges and opportunities. *Proceedings of Machine Learning and Systems*, 4, 795-813. <https://doi.org/10.5555/3520312.3520365>
- [3] Adepoju, O. E., Adejumo, A. O., & Oyeleye, B. A. (2023). Readiness of Nigerian teaching hospitals for artificial intelligence deployment: Infrastructure assessment and policy implications. *African Journal of Health Policy and Systems*, 12(2), 145-162. <https://doi.org/10.4314/ajhps.v12i2.8>
- [4] Nwosu, C. E., & Okereke, C. (2022). Energy infrastructure challenges in Nigerian teaching hospitals: Implications for digital health transformation. *West African Journal of Medicine*, 39(8), 897-904. <https://doi.org/10.55891/wajm.v39i8.1567>
- [5] García-Martín, E., Rodrigues, C. F., Riley, G., & Grahn, H. (2021). Estimation of energy consumption in machine learning. *Journal of Parallel and Distributed Computing*, 134, 75-88. <https://doi.org/10.1016/j.jpdc.2019.07.007>
- [6] Li, C., Tao, Y., Ao, W., Yang, S., & Bai, Y. (2023). Improving forecasting accuracy of daily enterprise electricity consumption using a random forest based on ensemble empirical mode decomposition. *Energy*, 165, 1220-1227. <https://doi.org/10.1016/j.energy.2018.10.113>
- [7] Dodge, J., Prewitt, T., Tachet des Combes, R., Odmark, E., Schwartz, R., Strubell, E., Luccioni, A. S., Smith, N. A., DeCario, N., & Buchanan, W. (2022). Measuring the carbon intensity of AI in cloud instances. *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, 1877-1894. <https://doi.org/10.1145/3531146.3533234>
- [8] Henderson, P., Hu, J., Romoff, J., Brunskill, E., Jurafsky, D., & Pineau, J. (2020). Towards the systematic reporting of the energy and carbon footprints of machine learning. *Journal of Machine Learning Research*, 21(248), 1-43. <https://doi.org/10.5555/3455716.3455964>
- [9] Strubell, E., Ganesh, A., & McCallum, A. (2020). Energy and policy considerations for modern deep learning research. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(9), 13693-13696. <https://doi.org/10.1609/aaai.v34i09.7123>
- [10] Patterson, D., Gonzalez, J., Le, Q., Liang, C., Munguia, L. M., Rothchild, D., So, D., Texier, M., & Dean, J. (2022). Carbon emissions and large neural network training. *arXiv preprint arXiv:2104.10350*. <https://doi.org/10.48550/arXiv.2104.10350>
- [11] Topol, E. J. (2023). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine*, 25(1), 44-56. <https://doi.org/10.1038/s41591-018-0300-7>

- [12] Chen, F., Taviani, V., Malkiel, I., Cheng, J. Y., Tamir, J. I., Shaikh, J., Chang, S. T., Hardy, C. J., Pauly, J. M., & Vasanawala, S. S. (2023). Variable-density single-shot fast spin-echo MRI with deep learning reconstruction. *Magnetic Resonance in Medicine*, 89(4), 1450-1466. <https://doi.org/10.1002/mrm.29521>
- [13] Campanella, G., Hanna, M. G., Geneslaw, L., Miraflor, A., Silva, V. W. K., Busam, K. J., Brogi, E., Reuter, V. E., Klimstra, D. S., & Fuchs, T. J. (2020). Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nature Medicine*, 25(8), 1301-1309. <https://doi.org/10.1038/s41591-019-0508-1>
- [14] Sendak, M. P., Gao, M., Brajer, N., & Balu, S. (2020). Presenting machine learning model information to clinical end users with model facts labels. *NPJ Digital Medicine*, 3(1), 41. <https://doi.org/10.1038/s41746-020-0253-3>
- [15] Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: Forecasting and control* (5th ed.). John Wiley & Sons. <https://doi.org/10.1002/9781118619193>
- [16] Jiang, P., Liu, Z., Wang, J., & Zhang, L. (2021). Decomposition-selection-ensemble forecasting system for energy consumption forecasting based on multi-objective optimization and fuzzy decision. *Applied Energy*, 283, 116254. <https://doi.org/10.1016/j.apenergy.2020.116254>
- [17] Taylor, S. J., & Letham, B. (2018). Forecasting at scale. *The American Statistician*, 72(1), 37-45. <https://doi.org/10.1080/00031305.2017.1380080>
- [18] Kong, W., Dong, Z. Y., Jia, Y., Hill, D. J., Xu, Y., & Zhang, Y. (2019). Short-term residential load forecasting based on LSTM recurrent neural network. *IEEE Transactions on Smart Grid*, 10(1), 841-851. <https://doi.org/10.1109/TSG.2017.2753802>
- [19] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735-1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- [20] Hong, T., & Fan, S. (2016). Probabilistic electric load forecasting: A tutorial review. *International Journal of Forecasting*, 32(3), 914-938. <https://doi.org/10.1016/j.ijforecast.2015.11.011>
- [21] Talavera-Llames, R. L., Pérez-Chacón, R., Martínez-Ballesteros, M., Troncoso, A., & Martínez-Álvarez, F. (2020). A nearest neighbours-based algorithm for big time series data forecasting. *Information Sciences*, 510, 196-214. <https://doi.org/10.1016/j.ins.2019.09.024>
- [22] Yuce, B., Li, H., Rezgui, Y., Petri, I., Jayan, B., & Yang, C. (2021). Utilizing artificial neural network to predict energy consumption and thermal comfort level: An indoor swimming pool case study. *Energy and Buildings*, 80, 45-56. <https://doi.org/10.1016/j.enbuild.2014.05.052>
- [23] Deb, C., Zhang, F., Yang, J., Lee, S. E., & Shah, K. W. (2020). A review on time series forecasting techniques for building energy consumption. *Renewable and Sustainable Energy Reviews*, 74, 902-924. <https://doi.org/10.1016/j.rser.2017.02.085>
- [24] Olatomiwa, L., Mekhilef, S., Huda, A. S. N., & Sanusi, K. (2021). Techno-economic analysis of hybrid renewable energy system for rural healthcare facilities in Nigeria. *Renewable Energy*, 157, 1200-1212. <https://doi.org/10.1016/j.renene.2020.04.145>

- [25] Okpala, C. N., Eze, C. U., & Ugwu, E. I. (2022). Cooling energy requirements in tropical hospital buildings: A Nigerian case study. *Energy and Buildings*, 254, 111589. <https://doi.org/10.1016/j.enbuild.2021.111589>
- [26] University College Hospital Ibadan. (2024). About UCH: Bed capacity and facility overview [Official website]. <https://uch-ibadan.org.ng>
- [27] BVGH. (n.d.). University of Nigeria Teaching Hospital, Enugu: Clinical trial site profile. BIO Ventures for Global Health. <https://www.bvgh.org>
- [28] Climate-Data.org. (2024). Climate data for Lagos, Kaduna, Enugu, Port Harcourt and Awka, Nigeria [Long-term climate normals]. <https://en.climate-data.org>
- [29] Humanitarian Data Exchange. (2020). Nigeria health facilities [Data set]. OCHA Nigeria. <https://data.humdata.org/dataset/nigeria-health-facilities>