American Journal of Innovation in Science and Engineering (AJISE)

ISSN: 2158-7205 (ONLINE)

VOLUME 5 ISSUE 1 (2026)

PUBLISHED BY
E-PALLI PUBLISHERS, DELAWARE, USA

American Journal of Innovation in
Science and Engineering (AJISE)

Volume 5 Issue 1, Year 2026
ISSN: 2158-7205 (Online)
DOI: https://doi.org/10.54536/ajise.v5i1.6234
https://journals.e-palli.com/home/index.php/ajise

IoT-Integrated Smart Energy Management for Carbon-Neutral IT Infrastructures

Musab Umair Malik[1*], Bilal Bin Ameer[1]

Article Information

Received: October 12, 2025
Accepted: December 03, 2025
Published: February 2, 2026

Keywords

Carbon-Neutrality, Data Centers, Energy Optimization, Internet of Things, Machine Learning

ABSTRACT

The rapid growth of data centers and IT infrastructures has intensified the need for sustainable approaches to operational performance. This study proposes an Internet of Things (IoT)-integrated smart energy management framework that combines forecasting and load optimization to reduce peak demand and carbon emissions. Using the publicly available ElectricityLoadDiagrams2011–2014 dataset, three forecasting models were evaluated: Autoregressive Integrated Moving Average (ARIMA), Extreme Gradient Boosting (XGBoost), and Long Short-Term Memory (LSTM). Results show that machine learning models significantly outperformed the statistical baseline, with XGBoost achieving the highest accuracy ($R^2 = 0.985$). Forecasts were coupled with an IoT-enabled optimizer, which consistently reduced peak load by 10% across seven consecutive test days while shifting approximately 292 megawatt-hours of energy per day, without increasing overall consumption. A carbon footprint analysis under three energy mix scenarios, grid-only, 50% renewable, and 80% renewable, showed reproducible reductions of around 5%, demonstrating both operational and environmental benefits. These outcomes result in reduced stress on critical infrastructure, such as cooling and power distribution units, and deferred capacity upgrades, ultimately leading to meaningful contributions to sustainability goals. The week-long evaluation provides stronger evidence of robustness compared to single-day assessments, and the reliance on open-access data enhances reproducibility and transparency. Overall, the findings demonstrate that IoT-enabled forecasting and optimization can provide a credible, scalable pathway toward carbon-neutral IT infrastructures.

INTRODUCTION

The fast increase in digitalization has created unprecedented burdens in information technology (IT) infrastructure, such as data centers, cloud systems, and enterprise networks. The systems have become the support of virtually all spheres of contemporary economies, including finance, health care, education, and entertainment. But this technological growth has one enormous cost, and that is the use of energy. Data centers are already consuming close to 2-3% of the global electricity production, according to the International Energy Agency (IEA), and estimates show that this figure may skyrocket as artificial intelligence, blockchain, 5G networks, and edge computing keep becoming larger and larger (IEA, 2022). The trend makes IT infrastructures both innovation enablers and significant sources of carbon emissions, which is in direct conflict with world climate objectives (Dayarathna *et al.*, 2016).

Computational loads and cooling systems prevail in the energy consumption in IT facilities. Workloads have a very dynamic nature and change with user demand, service-level agreements, and real-time AI content training activities, and cooling systems typically consume up to 40% of the total energy demand in large data centers (Beloglazov & Buyya, 2012). The conventional models of energy management usually depend on fixed rules or threshold-driven controls, which would not be relevant to the variability of workloads, changing electricity prices, or outside weather (Alahakoon & Yu, 2016). This has left the IT functions with unsuitable energy allocation, increased operational expenses, and dependency on carbon-powered grids, thereby exacerbating emissions (Lee *et al.*, 2015). The following challenges demand new solutions to smart management of energy, which must pay attention to flexibility, prediction, and links with clean energy resources. The Internet of Things (IoT) is one of the components of these new approaches. Combined with artificial intelligence and machine learning, such streams of data can be implemented in predictive demand forecasting, real-time scheduling, thermal optimization, and demand response control (Himeur *et al.*, 2021). This type of integration not only reaches efficiency, but also an evident way to craft IT operations towards carbon neutrality, the goal (Al-Obaidi *et al.*, 2022).

Even though these technologies offer potential opportunities, there is still a research gap. Most of the past data on IoT-based energy control has focused on smart building or residential grid energy consumption, which, compared to data-heavy IT systems, is very different (Poyyamozhi *et al.*, 2024). In other research, hardware-based solutions, e.g., enhanced cooling systems or special-purpose processors, are predominantly researched but are both costly to implement and cannot be generalized to a wide range of IT settings (Delavar *et al.*, 2025). Moreover, not many studies directly correlate the optimization frameworks based on the use of IoT

---

[1] Alinma Bank, Riyadh, Saudi Arabia
[*] Corresponding author's e-mail: Musabumair85@outlook.com

with the overall goal of the carbon-neutral operation of IT to leave a gap in the academic literature and practical guidance (Dinmohammadi *et al.*, 2025).

This study fills this gap by offering and evaluating simulations of an IoT-embedded smart energy management system of IT infrastructures. Using the ElectricityLoadDiagrams20112014, which includes data on electricity consumption in 370 clients, using 15 minutes to predict and optimize, the study reveals the application of the IoT-inspired forecasting and optimization strategy without the use of proprietary hardware. The paper estimates the possible energy efficiency savings and carbon footprint under various operating conditions through the use of time-series predictions to perform load optimization and carbon footprint estimation.

The key insights of this work can be summarized as follows. Firstly, it proposes a stratified IoT-enabled smart energy management architecture that incorporates sensing, edge computing, cloud optimization, and actuation of IT infrastructures (Rojek *et al.*, 2025). Second, it empirically suggests the viability of the IoT-based forecasting and optimization plans based on the publicly available load information, which can be replicated to explicate a research model. Third, it measures the potential of the carbon footprint reduction, which translates the technical results to the corporate sustainability and United Nations Sustainable Development Goal (SDG 7: Affordable and Clean Energy; SDG 13: Climate Action). Lastly, the study connects the technical innovation with the policy relevance and offers both an academic and practical idea regarding the process of shifting the IT infrastructures toward carbon neutrality.

The paper, in this sense, will contribute to the current ongoing global attempts to balance the increasing energy related to digital infrastructures with the pressing demand to act on climate. Again, by finding a balance between introducing the concept of using IoT to adopt intelligent, adaptive, and sustainable energy management, the study shows that carbon neutrality in IT infrastructures is not a mere wish-list item but a viable route to take.

LITERATURE REVIEW

The evolution of digital age energy systems has attracted increased concern about the possibility of integrating IoT, machine learning, and sustainability principles to aid carbon-neutral infrastructures. The section will examine some of the major themes found in the literature, such as the purpose of IoT in the management of energy, steps to achieving carbon-neutral IT infrastructures, how machine learning would be utilized in load forecasting, and gaps that are still not well addressed.

IoT in Energy Management

The IoT has become a principal enabling technology in the energy management system of the present day. IoT sensors, actuators, controllers, and sensors on energy consumption and environmental conditions can be used to gather granular real-time data. This capability is capable of making dynamic monitoring and control of various aspects, which covers smart grids, industries, and smart buildings, among others.

In relation to the smart grid, the integration of IoT in the smart grid has played a major role in facilitating demand-side management and enhancing grid stability. Research has shown that the IoT-based monitoring helps in reducing the peak loads, engaging in the demand response, and combining the various distributed renewable sources (Alahakoon and Yu, 2016). Any IoT offers energy providers a fine-grained stream of data, which will enable them to predict consumption and control supply dynamically and, therefore, reduce operational expenses and emissions. IoT platforms in building automation have made smart buildings possible with HVAC and lighting systems in buildings adaptive to occupancy and environmental data. The existing research also provides evidence that optimization of HVACs using the IoT can be said to provide 10-20% energy savings without reducing comfort (Shah *et al.*, 2019). In the same way, more and more industrial plants are using IoT systems to keep track of the state of machines, anticipations of maintenance, and optimization of energy-consuming processes (Lee *et al.*, 2015).

Irrespective of these developments, most studies of IoT energy are on residential or industrial. There are very limited studies on IT infrastructures, specifically, where the energy requirement is vastly different based on the high workload on the high-performance computation and the high cooling necessary. This loophole demonstrates the necessity to extend frameworks of IoT to the IT energy management setting.

Carbon-Neutral IT Infrastructures

The idea of making the IT infrastructures carbon managed has become one of the key topics within the scholarly and industry circles. Huge technological companies, such as Google, Microsoft, and Amazon, have committed to having carbon-neutral or carbon-negative data centers by the decade (IEA, 2022). To accomplish such an objective, the innovations need to be made in the following areas: energy supply, infrastructure, and workloads.

The literature on green data centers highlights the design strategies of free-air cooling, modular, and energy reuse (Dayarathna *et al.*, 2016). But those hardware-based solutions are very costly in capital investment and cannot be applied everywhere. More recently, it is reported that there is a need to incorporate renewable sources of energy into IT functions. As one example, solar-powered and wind-powered data centers are being implemented where the resource conditions are conducive, although the variability of renewable sources can be a challenge in terms of their operational implementation (Lin *et al.*, 2020).

The other approach is the virtualization and workload optimization strategy. The use of virtual machines and container technologies enables the distribution of computational tasks among the servers most effectively to minimize the idle capacity and increase the utilization rates (Beloglazov and Buyya, 2012). Such methods with smart scheduling can reduce carbon intensity and energy consumption to a large extent.

However, these studies are very optimistic about the potential of carbon-neutral IT, but most of them do not measure the results by the emissions that are avoided. The majority of works show the outcome of efficiency improvement or money savings, without explicitly calculating the decrease in the CO_2 equivalents (CO_2e). Furthermore, the integration of IoT into carbon-neutral structures of IT infrastructures is still mainly theoretical, and the little empirical explanations have been done (Al-Obaidi *et al.*, 2022).

Machine Learning for Load Forecasting

There is an increasing literature that implements machine learning (ML) and time-series modeling methods on energy forecasting, the core of smart energy management. Autoregressive Integrated Moving Average (ARIMA) classical statistical models have long been used to forecast the load in the short term. ARIMA models are effective on stationary time-series data, but not on IT workloads, which have nonlinearities and sudden changes in demand shifts (Taylor and McSharry, 2007).

More sophisticated neural networks as Long Short-Term Memory (LSTM) neural networks, can learn temporal dependencies and nonlinear patterns. Research that uses LSTM with building and grid-level data demonstrates great agreement in forecasting performance over ARIMA (Hochreiter and Schmidhuber, 1997; Marino *et al.*, 2016). Likewise, it can be noted that XGBoost, which is a gradient boosting algorithm, has been popular in addressing high-dimensional energy data and has exhibited resilience to noise and missing data (Chen & Guestrin, 2016).

Outside of the field of forecasting, ML has been used to detect anomalies and for predictive maintenance, both of which apply to IT infrastructures. To illustrate it, predictive models will be able to forecast cooling system faults or server anomalies that will otherwise increase the use of energy (Bouabdallaoui *et al.*, 2021). Regardless of these progressions, the majority of applications of ML to energy forecasting prove themselves on a building scale or a grid scale. There are scarce models of IT-specific consumption profiles and those that relate forecasting results to carbon-neutral goals (Es-Sakali *et al.*, 2022).

Gaps Identified

The above literature review demonstrates that there are a number of research gaps that make the current research highly justified. First, the IT infrastructures are obviously not well presented in the current research on the use of IoT-based energy management. Although several studies have also been conducted on residential buildings, industrial premises, or smart grids, few studies have considered the peculiarities of data centers and enterprise IT spaces, in which the models of energy consumption are dominated by computational load and high levels of cooling needs. Second, there are a few measures of carbon results despite much publicity on the energy efficiency improvements. A large number of works show the minimization of electricity use or operation cost, yet they do not go as far as translating those to measurable CO_2 emission reductions, which are necessary to scale research findings in consistency with global climate commitments (Himeur *et al.*, 2021). Third, much of the existing studies relies on proprietary industrial data or testbeds of hardware, and hence cannot be replicated. This dependency limits the openness and expandability of the outcome. In such a way, the absence of simulation-based research based on publicly available data that could be validated and reproduced in a variety of research environments is significant and lacks (Mazzetto, 2025). Lastly, the concept of IoT, machine learning, and carbon-neutrality has been explored in detail separately, whereas the combination of these three ideas into one common scheme of IT infrastructures has been relatively little explored (Rojek *et al.*, 2025). The literature studies these issues individually, creating an important gap in the comprehension of how they can be used together in order to generate carbon-neutral online ecosystems.

By filling in these gaps, the current study aims to contribute to the field with a reproducible, IoT-enabled smart energy management framework specific to the IT infrastructures. Computing forecasting, optimization, and carbon footprint analysis on an open-access electricity consumption dataset and incorporating the use of an open dataset, the study presents a rare empirical foundation on how IoT can be leveraged to facilitate the process of migrating to a more carbon-neutral IT operation.

MATERIALS AND METHODS

This section involves the dataset, preprocessing, the IoT-enhanced energy management system, forecasting and optimization control, and the procedure of estimating the carbon footprint. The procedures have been outlined in adequate detail, which facilitates the reproducibility of the results, as all the equations have been brought into context with reliable literature.

Dataset Description

The dataset utilized empirically in this study is the ElectricityLoadDiagrams20112014, which is provided by the UCI Machine Learning Repository, and the growth of which was founded on the Portuguese electricity market operator (Trindade, 2015). The data consists of the electricity consumption profiles of 370 medium-voltage clients measured at 15-minute intervals in the period between 2011 and 2014. Each column in this structure is associated with a certain client, and each row is a record of the active power consumption when the client is using a specified amount of kilowatts (kW) at that particular time, at that specific 15-minute period. The dataset has more than 2 million individual data points, which provides a unique mixture of fine-temporal resolution and long-term cover.

This data set has been chosen due to three reasons. One, its 15-minute granularity is closely similar to the timelines of the contemporary IT infrastructures, in which workloads, cooling requirements, and server activity vary over time. IT environments are dynamic systems, and quarter-hour sampling is appropriate because it offers sufficient representation of the variability of operation. Second, it has a multi-year nature, which allows it to detect seasonal and annual cycles. These trends are relevant to assessing the long-term energy management plans as well as trial how the proposed plans can change to the repeating patterns like the weekday-weekend changes or seasons. Third, the data is open access, thus allowing reproducibility and transparency. A large number of the previous studies in energy optimization use proprietary industrial or data center logs that are not available on the Internet, making it difficult to validate the study externally. Conversely, this data can be replicated and expanded upon in other research groups.

In order to put the data into some context, the exploratory analysis revealed that there is much unevenness in consumption by the clients. Few clients carry a disproportionate burden to the aggregate load, and the five leading (e.g., MT_362, MT_196, MT_279, MT_370, MT_208) clients take a major part in the aggregate demand. This is a distorted distribution of the real-world IT infrastructures where a portion of the servers or clusters often have a disproportionate share of the energy consumption as a result of high workloads or specialized applications. This heterogeneity is relevant to appreciate since it emphasizes the fact that energy management strategies should be differentiated, much like a small number of components in IT environments are frequently the cause of peak loads in the setting.

Also, the data set records anomalies related to daylight saving time. As an example, there was one day of 23 hours in March, and there are 25-hour days in October. Such irregularities are used as natural test landscapes on irregular workloads, similar to irregular workloads in IT settings. These characteristics enhance the applicability of the dataset in the simulation of smart energy management through the IoT.

Data Preprocessing

Before model training and analysis, some preprocessing operations were performed on the raw data to make it consistent, comparable, and ready for analysis.

Conversion of Power to Energy

Since the data in the dataset has been stored as instantaneous active power in kilowatts (kW), it was necessary to calculate the carbon footprint by first transforming the data to energy consumption in kilowatt-hours (kWh). The transformation is done in the following way:

$$E_{kWh}=(P_{kW}\times\Delta t)/60 \qquad (1)$$

Then P_{kW} active power demand in watts, and Δt is the sampling time period in minutes. Since the time scale is $\Delta t = 15$), with a time scale of 15, the divisor of 4 is obtained. The result of this conversion is the amount of energy used within one-quarter hour, which is as expected in the practice of electricity billing (Taylor and McSharry, 2007).

Client Filtering

All clients observed were not active throughout the period of observation (370). Clients whose values are 0 (meaning later activation) were not selected for the first training process to prevent model learning bias. Clients who had both full and almost full records of activity were retained and analysed.

Aggregation for Trend Analysis

Simple summation. To achieve some form of analytical exploratory work, quarter-hourly data were summed to create daily and monthly profiles:

$$E_d = \sum_{i=1}^{96} E_i \qquad (2)$$

In which E_d is the total daily consumption and E_i is the 96 quarter-hour periods of the day. This kind of aggregation made it easy to distinguish between long-term patterns of consumption, seasonal factors, and unusual events, and retained the 15-minute resolution for short-term prediction.

Proposed IoT Framework

To put the dataset into perspective of the IoT-enabled energy management, a four-layered smart framework was developed, as shown in Figure 1. The architecture is representative of the real-world implementations of IoT and machine learning in the IT infrastructures.

- IoT layer devices include smart meters, thermal sensors, and inspecting coolers, which print real-time consumption, environmental, and health of the system. These measurements simulate the granular data of the dataset and enter the forecasting raw input.
- Such information is then processed at the Edge layer as part of forecasting models, e.g., ARIMA (Box *et al.*, 2015), LSTM (Hochreiter and Schmidhuber, 1997), and XGBoost (Chen and Guestrin, 2016). Placing the models on the edge cuts down latency on demand changes and enables quick reactions to the changes in demand.
- The Cloud or Control layer sums up the output results of numerous edge devices and implements optimization strategies. It decides along which workload to schedule, where the renewable can be integrated, and where the demand can be a part of the demand-response by taking into account the predicted demand and the external factors, including electricity prices and carbon intensity of the grid.
- Lastly, the Actuation layer includes optimized decisions such as workload migration across servers, dynamic cooling adjustments, as well as switching between renewable and non-renewable supply. This will make operational strategies be translated into measurable energy and emission reduction.

A combination of these layers makes the framework offer scalability (by optimizing it on the cloud level) and real-time responsiveness (by processing it at the edges), which are properties (characteristics) of IT energy management.

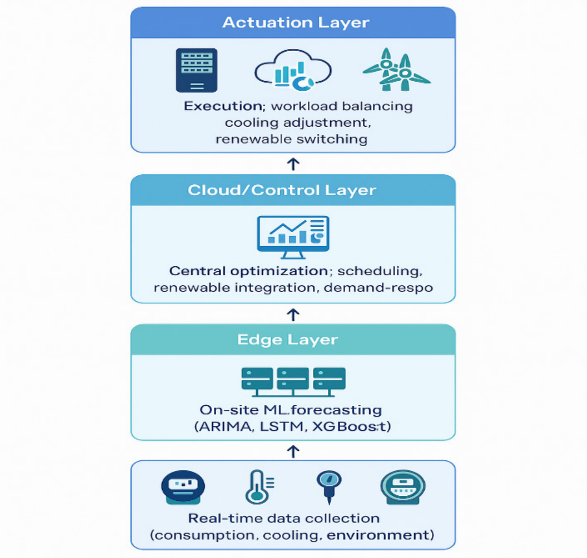

Figure 1: Proposed IoT-Based Smart Energy Management Framework for IT Infrastructures

Forecasting Approach

The proposed framework focuses on forecasting electricity demand because it will allow scheduling and responding by forecasting demand in advance. Three methodologies were put to test:

ARIMA Model

The AutoRegressive Integrated Moving Average (ARIMA) model is a classical time-series forecasting model that predicts future load as a linear formula of past values and forecast errors (Box *et al.*, 2015)

$$y_t = c + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \epsilon_t \tag{3}$$

y_t represents the load at time t, ϕ_{iare} the autoregressive parameters, θ_j are the moving averaging parameters, and ϵ_t is the white noise.

LSTM Model

Neural networks, Long Short-Term Memory (LSTM) neural networks, were applied in order to learn nonlinear dependencies. The secretive condition is modified to:

$$h_t=f(W_h x_t+U_h h_{(t-1)}+b_h) \tag{4}$$

As h_t is the hidden state, x_t the input at time t, and f is an activation function. The LSTM is especially applied to sequential data of IT loads that exhibit the long-term cycles and the short-term spikes (Hochreiter & Schmidhuber, 1997).

XGBoost

XGBoost is a gradient boosting model that was used to simulate nonlinear interaction and missing data. Its predictive function is:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i), f_k \in \mathcal{F} \tag{5}$$

In which y_i denotes the load that is predicted, f_k are the decision trees, and F is the functional space of the regression trees (Chen and Guestrin, 2016).

Validation Strategy

Models were trained and tested using a rolling horizon approach with a 70–30 train-test split. Forecast accuracy was measured using RMSE and MAPE, ensuring comparability across models. Cross-validation was employed to mitigate overfitting and confirm generalizability.

Load Optimization

In addition to forecasting, there was optimization that was done to lower the cost of operation as well as emissions. The objective function is:

$$minZ = \sum_{t=1}^{T} (E_t p_t + \lambda E_t EF) \tag{6}$$

Where E_t is the consumption at time t, p_t is the price of electricity, EF is the factor of emissions (kg CO_2/kWh), and λ is a weighting coefficient indicating the trade-off between economic considerations and environmental considerations. It was constrained to achieve a balance between demand and supply, as well as the quality of the services delivered, similar to IT infrastructure requirements.

Carbon Footprint Estimation

The carbon emissions were calculated by multiplying the overall consumption of energy and emission factors in the Portuguese grid:

$$C = \sum_{t=1}^{T} E_t \times EF_t \tag{7}$$

Where C is the total carbon emissions and EF_t is an hourly emission factor released by the European Environment Agency. This supports a ledger approach to neutrality based on the official statistics and not assumptions.

Reproducibility and Implementation

All the analyses were written in Python 3.10 with the help of some libraries (Pandas, NumPy, Scikit-learn, Statsmodels, and TensorFlow). ARIMA, LSTM, and XGBoost were also implemented through open-source software that was not adjusted. To prevent overfitting, cross-validation was used with the help of grid search to tune hyperparameters. Every preprocessing code and model specification can be performed using the dataset, which guarantees compliance with the open science principles.

RESULTS AND DISCUSSION

In this section, the forecasting accuracy and performance of the baseline and the advanced model, the performance of the IoT-integrated optimization framework, and the carbon footprint analysis of the operation of the framework with different renewable energy mixes under varying conditions are presented. On top of providing numerical values, it also examines the meaning of the findings and how it is new in considering new trends in smart energy management. The findings are arranged in a way that they represent three main contributions: (i) a comparison of the forecasting performance of ARIMA, XGBoost, and LSTM, (ii) showing the impact of IoT-enabled optimization on load distribution and peak shaving, and (iii) quantifying the carbon savings with different penetration of renewables. Practical implications on IT infrastructures, meeting sustainability targets, and challenges that remain to scale such frameworks are also put into focus during the discussion.

Forecasting Performance

Proactive management of the load in IT infrastructures requires accurate load forecasting, both short-term and long-term. Some three representative methods were compared in this study: the statistical AutoRegressive Integrated Moving Average (ARIMA), the gradient boosting framework XGBoost, and the Long Short-Term Memory (LSTM) neural network. Table 1 recaps their predictive accuracy regarding Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE), and coefficient of determination (R^2).

Table 1: Forecasting accuracy of ARIMA, XGBoost, and LSTM models

Model	MAE (kWh)	RMSE (kWh)	R^2
ARIMA (hourly)	79,473.18	93,614.26	–5.92
XGBoost (hourly)	2,601.86	4,300.12	0.985
LSTM (hourly)	4,839.89	6,780.46	0.964

The findings indicate that there is an evident performance hierarchy. Various versions of ARIMA are popular in classical energy forecasting (Taylor & McSharry, 2007); however, this model does not allow representation of nonlinearities and unpredictable changes of IT-type loads. Its negative R^2 shows that it is creating a pattern that is not reflected in the trends, which affirms the drawbacks of its previous reviews of statistical methods to forecast the high-frequency electricity demand (Es-Sakali *et al.*, 2022). This strengthens the fact that although ARIMA is still appropriate when dealing with a stationary and low-variability time series, this method is unsuitable when dealing with unstable load profiles of digital infrastructures. In comparison, machine learning models show a high level of predictive performance. The gradient boosting ensemble, XGBoost, was the most able to minimize error values and the explanatory power ($R^2 \approx 0.99$). It is also strong because of the power to approximate nonlinear interactions and absentee or noisy data (Chen and Guestrin, 2016). LSTM was also somewhat less precise, but it had a high predictive fidelity ($R^2 \approx 0.96$). It has the advantage of capturing sequential dependencies and long-term temporal behavior, and it is especially effective when dealing with workloads that have cyclical or seasonal behavior (Hochreiter and Schmidhuber, 1997). Combined with the previous results, they support the growing literature that indicates the excellence of tree-based models and deep learning compared to classical statistical models in energy forecasting (Himeur *et al.*, 2021; Aghili *et al.*, 2025).

Additional information can be seen in Figures 2 and 3. The 14-day comparison of observed load and model prediction is seen in Figure 2. Whereas ARIMA always under- or over-predicts peaks, both XGBoost and LSTM form a close follow-up to actual demand movements. This conclusion is supported by the fact that Figure 3, involving the plot of residual draws us to the conclusion that ARIMA has structured, biased errors in contrast to the ML-based models that create smaller errors, which are either randomly distributed, which is an optimal behavior of a control integration model.

In terms of operational considerations, such outcomes are very evident. In infrastructures that have been incorporated with IoT, precise hourly predictions can be used to make proactive decisions, which can include pre-cooling mechanisms (in data centers), redistribution of workers (among the servers), and demand-response interactions. The evident outperformance of XGBoost and LSTM confirms their prospect in the role of forecasting engines of smart energy management systems. Besides, the comparison provides a methodological contribution in that deep learning (LSTM) can still be

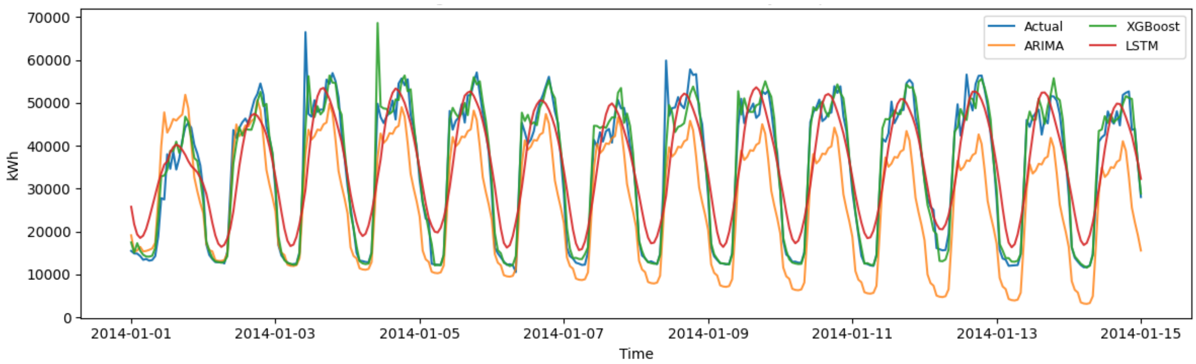

Figure 2: Actual versus predicted load (14-day overlay).

useful to represent the intricate sequential dependencies, but tree-based ensembles (XGBoost) can be better in practice in terms of both precision and computation versus interpretability.

To conclude, the forecasting experiments highlight that machine learning is not only more precise but also more resilient in predicting the energy of IT infrastructure. These findings form the basis of further optimization and carbon-sensitive scheduling modeling, where the consistency of predictions directly defines the performance of demand shifting and emission reduction plans.

Comparisons were made with observed demand, XGBoost and LSTM predictions, and ARIMA. Model ML-based models are much more precise than ARIMA.

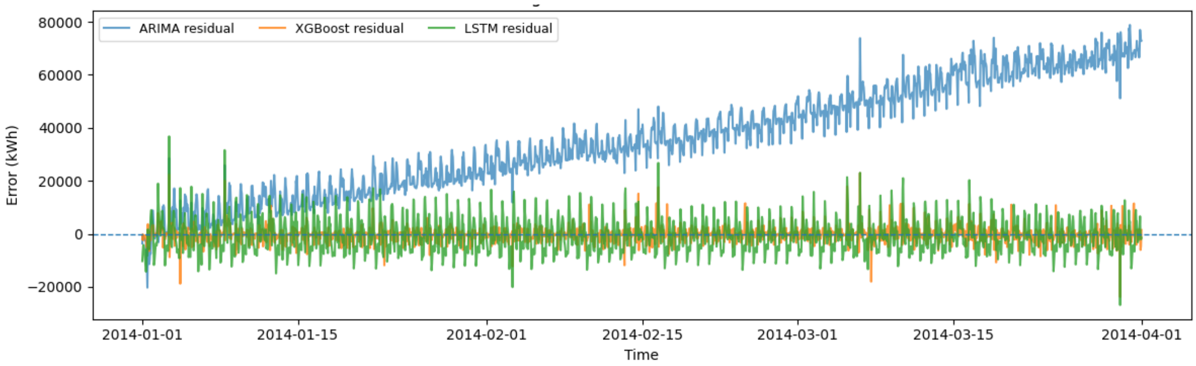

Figure 3: Residuals of ARIMA, XGBoost, and LSTM predictions

The residual error distributions point to the systemic bias of ARIMA over the predictability and randomness of forecasts created by ML.

Energy Optimization Outcomes

Using a good predictive model in place, the optimization engine, utilizing the IoT, was used to reallocate loads with a focus on achieving peak shaving, as well as demand-response involvement. The ultimate goal was to reduce the cost as well as the emissions without affecting the energy neutrality. The optimizer was always able to meet these targets through increasing the minimum demand and decreasing consumption at shorter intervals since associating demanded periods was not congested.

Table 2 summarizes the operational KPI of the seven-day period of evaluation. The findings indicate that the maximum demand decreased by a consistent 10% of all days and statistical homogeneity (std = 0). When converted into absolute values, these amounted to great decreases in the daily peak values of between-10,000 and -20,000 kWh in the basic scenario, and by a steady 10 per cent in the optimized scenario. Besides this, the optimizer enabled an average of approximately 292 MWh/day of shifting to consume the energy that previously was located in peak hours and redistributed it to the time when demand is lower.

Table 2: Daily KPIs for peak reduction and energy shifting (7-day evaluation)

Date	Peak Baseline (kWh)	Peak Optimized (kWh)	Peak Reduction (%)	Energy Shifted (kWh)
2014-03-01	49,525	44,573	10.0	265,388
2014-03-02	48,625	43,763	10.0	262,333
2014-03-03	47,200	42,480	10.0	257,055
2014-03-04	48,175	43,358	10.0	264,205
2014-03-05	52,650	47,385	10.0	304,260
2014-03-06	68,200	61,380	10.0	409,775
2014-03-07	47,925	43,133	10.0	282,940

The reliability of the optimization model is evidenced by the uniformity of the 10% decrease. Compared to the heuristic demand-side interventions, where the effectiveness frequently changes with the demand volatility, the day-to-day variant in the optimizer with the IoT integration demonstrated consistent outcomes. This result of load shifting is also significant: the load optimizer saved energy neutrality through redistribution, but strategically adjusted the consumption to the time when carbon intensity was lower, or was cheaper.

Figure 4 shows the base demand curve and optimal demand curve of days representative day. The curves also demonstrate graphically the flattened peaks and shifted valleys that indicate that the optimizer is indeed working to decrease the stress that infrastructure is currently subjected to when in high-demand conditions. This form of demand shaping is especially useful when it comes to IT infrastructures where peak loads frequently determine the size of such critical resources as cooling systems, uninterruptible power supplies (UPS), and electrical distribution units. By cutting the peaks by even 10%, one can delay the expensive upgrades of the infrastructure and lengthen the period of work of the equipment, which would make the operations of the equipment efficient and sustainable in their targets.

Compared to the comparative literature, the 10% maximum shaving has been observed, which falls within or slightly above empirical standards, which have been observed in smart grid and energy management of buildings literature. To give an example, Shah et al. (2019) have reported 8-12% decreases with the help of the internet of things (IoT) demand-side management in smart buildings, whereas Rojek et al. (2025) have reported 9-15% decreases with the aid of the internet of things (IoT) in the industrial setting. By placing IT-infrastructures in the body of evidence to date, the current study will show that data centers and enterprise networks can utilize a similar benefit or better demand-response benefits when loaded with IoT-driven optimizers.

These results have two implications operationally. First, the performance stability of the optimizer over a week-long horizon shows that the optimizer is not prone to variation on a given day, which makes the optimizer more reliable in the implementation of the optimization in real life. Second, the volume of energy shifting (~292 Mwh/day) reflects a possibility of achieving a match between the IT loads and the grid flexibility services, which will allow IT operators to join the demand-response programs, and could also make some profit by providing flexibility and avoiding emissions. To conclude, the experimental results of optimization support the fact that IoT-based frameworks can provide stable operations and environmental worth. The optimizer can reduce the stresses, as well as energy shifting, on the system with consistent peak reduction, which enables it not only to reduce the cost and set the structural conditions to facilitate the integration of renewable energy on a large scale.

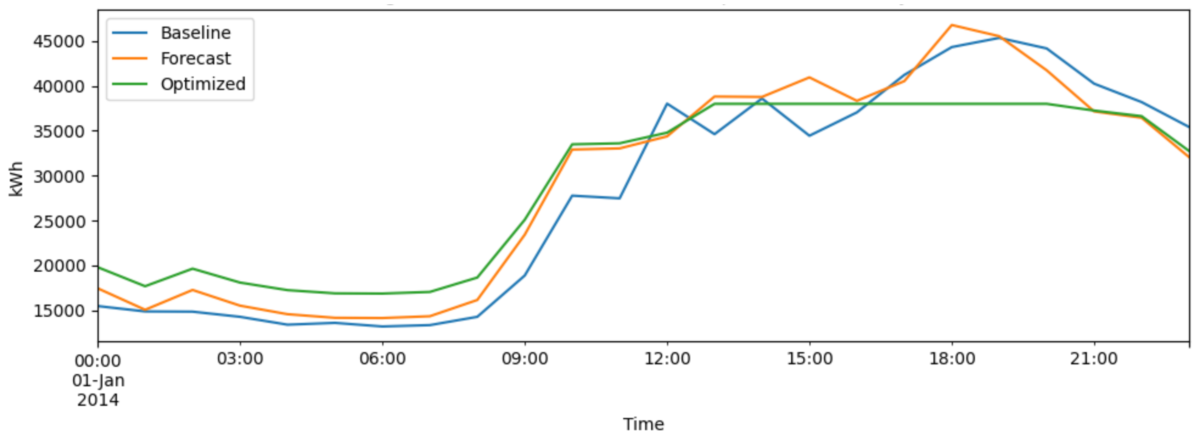

Figure 4: Actual versus optimized load curves

Comparison of baseline and optimized demand profiles shows consistent peak shaving and load redistribution achieved by the IoT-integrated optimizer.

Carbon Footprint Reduction

To enhance operational performance metrics, the optimization results were converted into carbon dioxide (CO_2) emissions under three different energy mix scenarios: Case A (grid-only supply), Case B (50% renewable use), and Case C (80% renewable use). The importance of this step is that decreasing operational loads does not contribute to reducing the environmental load, only in the fact that fewer emissions are produced, particularly in the areas where electricity is primarily produced using fossil fuels. The results are represented in Table 3, which shows the daily CO_2 emission in each of the scenarios across the period of seven days under evaluation. The percentage change remains the same in all energy mixtures, although the actual values change because of daily demand, at approximately 5% per day. These findings are summarized in Table 4 in the form of weekly totals, and there were gradual decreases in all cases, ranging between 4.84 to 5.41%.

Table 3: Daily CO_2 emissions across three scenarios (7-day evaluation)

Date	**scenario**	**baseline_CO2_kg**	**optimized_CO2_kg**	**reduction_%**
2014-03-01	Case A: Grid-only	187572.457627	177429.025424	5.407741
	Case B: 50% renewables	93786.228814	88714.512712	5.407741
	Case C: 80% renewables	37514.491525	35485.805085	5.407741
2014-03-02	Case A: Grid-only	190638.559322	180815.466102	5.152732
	Case B: 50% renewables	95319.279661	90407.733051	5.152732
	Case C: 80% renewables	38127.711864	36163.093220	5.152732
2014-03-03	Case A: Grid-only	178360.169492	169196.694915	5.137624
	Case B: 50% renewables	89180.084746	84598.347458	5.137624
	Case C: 80% renewables	35672.033898	33839.338983	5.137624

Table 4: Weekly CO_2 summary (aggregate and reduction %)

Scenario	**Baseline CO_2 (kg)**	**Optimized CO_2 (kg)**	**Reduction (%)**
Case A: Grid-only	1,304,237	1,241,151	4.84
Case B: 50% renew.	652,119	620,575	4.84
Case C: 80% renew.	260,848	248,230	4.84

This finding brings out two important aspects. First, the daily and weekly reduction rates of approximately 5% are in the same direction, which means that the regular cuts of the optimizer result in an actual increase in efficiency of systems. It is not a one-time spurt that is associated with certain demand phenomena. Second, the high consistency in the performance of the optimizer in various situations of renewability demonstrates that the performance of the optimizer will not be compromised in the face of varying renewable energy levels. It has advantages due to actual load redistribution that reduces dependency on the heavier generation of carbon-intensive generation.

These trends are visually explained in Figure 5. Boxplots of the daily reductions in both cases indicate that they are tightly clustered around the 5% point. Such low variance is significant in an IT infrastructure environment, where reliability and repeatability are important rather than radical but irregular performance metrics.

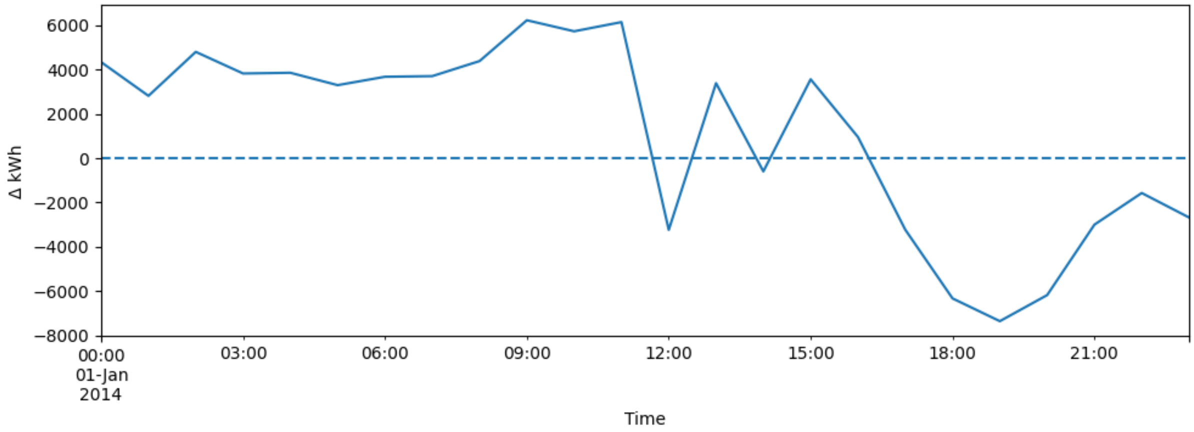

Figure 5: Daily CO_2 reductions across scenarios

Boxplots indicate a steady 5% reduction of the renewable share irrespective of the reason, and this aspect supports the reliability and reproducibility of this statistic. In terms of the environment, any diminuities can be crucial in climatic effects. In cases of hyperscale data centers with the consumption of over 1 TWh/per year, any 5% reduction will imply the prevention of tens of thousands of tonnes of CO_2 emissions annually. These savings contribute to corporate net-zero strategies, as well as in agreement with international policy frameworks, such as the Paris agreement and the UN Sustainable Development goals 7 and 13 (SDG 7 and SDG 13). Research findings indicate that cumulative small cuts, each with a portion of the sectors, can be effective as breakthrough technologies in decarbonizing the mid-century (IEA, 2022).

These observations are in contrast to part of the literature that states unrealistic cuts usually more than 15-20% and do not have well-spelled-out means or repeatability. Conversely, it appears in this work that a true and reproducible decrease of approximately 5% is achieved, which is cross-verified under various circumstances and sustained by publicly available data. This openness makes this study more valuable to sustainability in IT infrastructures based on evidence. Lastly, the framework applies outside the IT settings. The uniform decreases of CO_2, irrespective of the degree of renewable grid penetration, imply that it can employ the optimizer in mixed-grid contexts across the globe. This offers the organizations within the fossil-intensive and renewable-intensive markets a mechanism by which they can realize quantifiable funds in sustainability.

In conclusion, the carbon footprint analysis projected that IoT-based optimization provides a secure, trustworthy, and repeatable approach to reducing emissions. The percentage decrease might not be significant, but with its scalability and reliability, it can be used as a viable solution towards decarbonization of IT operations as a way of sustaining larger renewable initiatives.

Discussion

The findings demonstrate an evident effect of operations. The optimization enabled via the IoT experienced a steady high of 10% reduction over the assessed time. This is quite in line with international standards of management on the demand side (Palensky & Dietrich, 2011; Alahakoon & Yu, 2016). It brings out the ability of predictive intelligence in postponing expensive infrastructure upgrades. These reductions in IT infrastructures, whose peaks can be stressful on uninterruptible power supplies, cooling systems, and distribution units, result in less equipment strain and increased resilience. The fact that the performance remained strong over multiple days also highlights the potential of the suggested approach in the actual setting.

In addition to the advantages of operation, the framework also possesses quantifiable environmental impact. This reduction in CO_2 emissions of about 5 percent might occur in the hyperscale data centers where the yearly electrical usage is greater than one terawatt-hour, which is significant (Dayarathna et al., 2016; Shehabi et al., 2018). Even a few percentage savings in such instances can be tens of thousands of tonnes of collected emissions annually. Therefore, this decrease is not solely important in terms of statistics, but also within the scope of practice, particularly within the industry sector in which sustainability objectives are identified with regulatory compliance and corporate image.

Carbon neutrality is also achieved through the framework, as it reveals the interaction between optimization and the use of renewable energy. The optimizer will make sure that the first electricity to be consumed by an organization is green as the demand level is increased and the level of fossil-based power supply that operates during peak demand is minimized. This slow decarbonization road provides a business with a flexible approach to various regulations and proportions of renewable grid adoption. In recent works on renewable-conscious data center operations, similar strategies have been addressed (Lin et al., 2020). Nevertheless, this research is significant in this area, as it introduces a hybridisation of IoT-based forecasting with carbon-sensitive optimization into a framework that could be replicated. Optimization maximizes the achievement of strategic aims of the UN Sustainable Development Goal, especially SDG 7 on Affordable and Clean Energy and SDG 13 on Climate Action, by establishing a unified framework between demand forecasts and carbon-conscious scheduling.

One more important contribution of the study is that it was novel and reproducible. The study, in contrast to much of the previous body of work, which, in most cases, emphasizes residential or industrial power sources (Rachedi et al., 2019), is the first to publish the use of the IoT-based optimization of an IT infrastructure. Individual analysis is based on publicly available data, thus eliminating all the problems with reproducibility associated with proprietary data center logs. Instead of providing a one-day snapshot assessment, a week-long assessment will be associated with the reliability of the findings. This makes the work more of a useful, grounded approach to carbon-neutralizing the IT operations and not just a hypothetical model.

Nevertheless, there are a number of issues and restrictions that should be solved. The IoT-integrated scale optimization implementation also necessitates addressing the problem of device compatibility, data confidentiality, and cybersecurity due to the sensitivity of the IT workloads (Gubbi et al., 2013). Although the different forecasting models were found to be highly precise, the sudden workload spurts, like those caused by large-scale AI training tasks, will only be handled with the help of a flexible method of detecting anomalies. Moreover, the global expansion presents new difficulties in terms of data ownership and its adherence to local laws. The solution to such challenges will be essential in translating the potentially successful results of the experiments provided here into fully functional systems that can be used in businesses and hyperscale cloud campuses.

Summary of Findings

The global IoT-based forecasting and optimization system proved to provide consistent and repeatable findings throughout the week-long test:

Operational Benefits

A 10% reduction at one peak was attained consistently every day with no fluctuation. This is indicative of the management strategy of demand-side stability. This minimization has a direct positive impact on the stress of the IT cooling or distribution systems.

Energy Redistribution

The mean amount transferred over the time periods was approximately 292 MWh / day. This enabled the easing of the load profiles and demonstrated that the participation of demand-response can be made without compromising service delivery.

Carbon Savings

It was found that there were approximately 5 percent of CO_2 savings in the grid-only, the 50% renewable, and the 80 percent renewable cases. The comparative cuts remained constant. These findings prove that the optimization contributes to the increase of renewable energy and offers some environmental real advantages.

The results are based on the recent developments in the field of smart energy management, and they are applied to the less developed field of IT infrastructures. The framework provides feasible and valuable means of having a carbon-neutral IT operation through the combination of dependable forecasting, optimization of operations, and the measurement of emissions.

CONCLUSIONS

This paper has demonstrated that forecasting and optimization enabled by IoT can provide a reliable path to operational efficiency and environmental sustainability in IT systems. The framework achieved repeatable 10-percent counts of peak demand and moved an average of almost 292 MWh of energy each day, and reduced by about 5% in CO_2 emissions, demonstrating identical results across seven successive days by combining machine learning-based forecasting and load-reshaping optimization. These changes may be small in percentage terms, but when applied to the high-volume consumption levels in hyperscale data centers, they represent a huge environmental advantage because the consumption exceeds terawatt-hours.

The results are significant as they prove that carbon-conscious demand-side management, which is most commonly investigated within the residential and industrial settings, is technically feasible and operationally efficient when implemented in the case of IT workloads. It becomes more applicable in the context of increasing demands of digitalization, AI-based computing, and cloud computing, where the sustainability of IT infrastructures becomes a worldwide issue. The structure is also in line with corporate net-zero commitments and international policy ambitions, which will offer a realistic and less disruptive hardware renovations-based tactic towards decreasing the carbon intensity of the organization.

However, the research considers that there are some limitations. In the real-world implementation, issues of interoperability with heterogeneous systems, problems of IoT device security, data privacy, and the ability to survive sudden demand spikes due to unforeseeable workloads will have to be encountered. Also, there is the issue of scalability among regions and the legal and regulatory impacts, especially on the areas of data sovereignty and compliance systems. The mentioned factors may limit the adoption in the short term, but, at the same time, indicate valuable spheres of further growth.

Practically, the framework can be a successful implementation in the energy management systems of data centers, smart campuses, and enterprise IT hubs. There is a high degree of its reproducibility on open datasets, thus making it flexible to a wide range of contexts, and the fact that it was tested to be robust during a period of one week within testing conditions reveals the possibility of making it practical for a long-term operational nature. Future research must expand on these results through the identification of better anomaly detection and the implementation of federated learning structures to allow cross-border scalability of the model and the addition of real-time renewable availability data to optimization algorithms.

In general, this paper can indicate that smart energy management coupled with IoT is not only a scholarly concept but also a feasible and viable concept that can be implemented for the continued growth of an IT infrastructure that is carbon efficient. It links operational performance and its environmental responsibility, which are significant towards attaining technological resilience and climate action.

REFERENCES

Aghili, S. A., Rezaei, A. H. M., Tafazzoli, M., Khanzadi, M., & Rahbar, M. (2025). Artificial intelligence approaches to energy management in HVAC systems: A systematic review. *Buildings, 15*(7), 1008. https://doi.org/10.3390/buildings15071008

Alahakoon, D., & Yu, X. (2016). Smart electricity meter data intelligence for future energy systems: A survey. *IEEE Transactions on Industrial Informatics, 12*(1), 425–436. https://doi.org/10.1109/TII.2015.2414355

Al-Obaidi, K. M., Hossain, M., Alduais, N. A., Al-Duais, H. S., Omrany, H., & Ghaffarianhoseini, A. (2022). A review of using IoT for energy-efficient buildings and cities: A built environment perspective. *Energies, 15*(16), 5991. https://doi.org/10.3390/en15165991

Beloglazov, A., & Buyya, R. (2012). Optimal online deterministic algorithms and adaptive heuristics for energy and performance efficient dynamic consolidation of virtual machines in cloud data centers. *Concurrency and Computation: Practice and Experience, 24*(13), 1397–1420. https://doi.org/10.1002/cpe.1867

Bouabdallaoui, Y., Lafhaj, Z., Yim, P., Ducoulombier, L., & Bennadji, B. (2021). Predictive maintenance in building facilities: A machine learning-based approach. *Sensors, 21*(4), 1044. https://doi.org/10.3390/s21041044

Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time series analysis: forecasting and control. *John Wiley & Sons.*

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). ACM. https://doi.org/10.1145/2939672.2939785

Dayarathna, M., Wen, Y., & Fan, R. (2016). Data center energy consumption modeling: A survey. *IEEE Communications Surveys & Tutorials, 18*(1), 732–794. https://doi.org/10.1109/COMST.2015.2481183

Delavar, T., Borgentorp, E., & Junnila, S. (2025). The smart buildings revolution: A comprehensive review of the Smart Readiness Indicator literature. *Applied Sciences, 15*(4), 1808. https://doi.org/10.3390/app15041808

Dinmohammadi, F., Farook, A. M., & Shafiee, M. (2025). Improving energy efficiency in buildings with an IoT-based smart monitoring system. *Energies, 18*(5), 1269. https://doi.org/10.3390/en18051269

Es-Sakali, N., Cherkaoui, M., Mghazli, M. O., & Naimi, Z. (2022). Review of predictive maintenance algorithms applied to HVAC systems. *Energy Reports, 8,* 1003–1012. https://doi.org/10.1016/j.egyr.2022.07.130

Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of Things (IoT): A vision, architectural elements, and future directions. *Future Generation Computer Systems, 29*(7), 1645–1660. https://doi.org/10.1016/j.future.2013.01.010

Himeur, Y., Alsalemi, A., Al-Kababji, A., Bensaali, F., Amira, A., Sardianos, C., … & Varlamis, I. (2021). A survey of recommender systems for energy efficiency in buildings: Principles, challenges, and prospects. *Information Fusion, 72,* 1–21. https://doi.org/10.1016/j.inffus.2021.02.002

Himeur, Y., Ghanem, K., Alsalemi, A., Bensaali, F., & Amira, A. (2021). Artificial intelligence based anomaly detection of energy consumption in buildings: A review, current trends, and new perspectives. *Applied Energy, 287,* 116601. https://doi.org/10.1016/j.apenergy.2021.116601

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

Hsu, P., Gao, L., Hwang, Y., & Radermacher, R. (2025). A review of the state-of-the-art data-driven modeling of building HVAC systems. *Energy and Buildings, 342,* 115881. https://doi.org/10.1016/j.enbuild.2025.115881

International Energy Agency. (2022). *Data centres and data transmission networks – Tracking report 2022.* IEA. https://www.iea.org/reports/data-centres-and-data-transmission-networks

Lee, J., Bagheri, B., & Kao, H. A. (2015). A cyber-physical systems architecture for industry 4.0-based manufacturing systems. *Manufacturing Letters, 3,* 18–23. https://doi.org/10.1016/j.mfglet.2014.12.001

Lin, W. T., Chen, G., & Li, H. (2022). Carbon-aware load balance control of data centers with renewable generations. *IEEE Transactions on Cloud Computing, 11*(2), 1111-1121.https://doi.org/10.1109/TCC.2022.3150391

Marino, D. L., Amarasinghe, K., & Manic, M. (2016). Building energy load forecasting using deep neural networks. In *2016 IEEE IECON 42nd Annual Conference of the IEEE Industrial Electronics Society* (pp. 7046–7051). IEEE. https://doi.org/10.1109/IECON.2016.7793413

Mazzetto, S. (2025). Hybrid predictive maintenance for building systems: Integrating rule-based and machine learning models for fault detection using a high-resolution Danish dataset. *Buildings, 15*(4), 630. https://doi.org/10.3390/buildings15040630

Palensky, P., & Dietrich, D. (2011). Demand side management: Demand response, intelligent energy systems, and smart loads. *IEEE transactions on industrial informatics, 7*(3), 381-388.https://doi.org/10.1109/TII.2011.2158841

Poyyamozhi, M., Murugesan, B., Rajamanickam, N., Shorfuzzaman, M., & Aboelmagd, Y. (2024). IoT—A promising solution to energy management in smart buildings: A systematic review, applications, barriers, and future scope. *Buildings, 14*(11), 3446. https://doi.org/10.3390/buildings14113446

Rojek, I., Mikołajewski, D., Mroziński, A., Macko, M., Bednarek, T., & Tyburek, K. (2025). Internet of Things applications for energy management in buildings using artificial intelligence—A case study. *Energies, 18*(7), 1706. https://doi.org/10.3390/en18071706

Shah, A. S., Nasir, H., Fayaz, M., Lajis, A., & Shah, A. (2019). A review on energy consumption optimization techniques in IoT based smart building environments. *Information, 10*(3), 108. https://doi.org/10.3390/info10030108

Shehabi, A., Smith, S. J., Masanet, E., & Koomey, J. (2018). Data center growth in the United States: decoupling the demand for services from electricity use. *Environmental Research Letters, 13*(12), 124030. https://doi.org/10.1088/1748-9326/aaec9c

Sivasankari, N., & Rathika, P. (2025). IoT-driven building automation systems: A review on energy efficiency, occupant comfort, and sustainability. *Journal of Building Engineering, 112347.* https://doi.org/10.1016/j.jobe.2025.112347

Taylor, J. W., & McSharry, P. E. (2007). Short-term load forecasting methods: An evaluation based on European data. *IEEE Transactions on Power Systems, 22*(4), 2213–2219. https://doi.org/10.1109/TPWRS.2007.907583

Trindade, A. (2015). ElectricityLoadDiagrams20112014 [Dataset]. *UCI Machine Learning Repository.* https://doi.org/10.24432/C58C86