I.J. Modern Education and Computer Science, **2026, 5, 168-191**
Published Online on October 8, 2026 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijmecs.2026.05.10

Multi-Horizon Deep Learning Framework for Sustainable Energy Forecasting in Smart Campuses: Enabling Energy Efficient Smart Educational Buildings

Ajay Kumar
Department of Computer Science and Engineering, University Institute of Engineering and Technology, Maharshi Dayanand University, Rohtak, 124001, India
Email: ajay.rs.uiet@mdurohtak.ac.in
ORCID iD: https://orcid.org/0000-0002-3693-9701

Rainu Nandal*
Department of Computer Science and Engineering, University Institute of Engineering and Technology, Maharshi Dayanand University, Rohtak, 124001, India
Email: rainunandal.uiet@mdurohtak.ac.in
ORCID iD: https://orcid.org/0000-0003-0350-0388
*Corresponding author

Kamaldeep Joshi
Department of Computer Science and Engineering, University Institute of Engineering and Technology, Maharshi Dayanand University, Rohtak, 124001, India
Email: kduiet@mdurohtak.ac.in
ORCID iD: https://orcid.org/0000-0002-3238-0234

Received: October 10, 2025; Revised: December 20, 2025; Accepted: February 13, 2026; Published: October 8, 2026

Abstract: The accurate energy consumption forecast of educational institutions is critical for sustainable energy management and intelligent campus infrastructure. Classical forecasting models fail to simultaneously capture local temporal variations and long-range dependencies, thus diminishing their reliability across different time horizons. To overcome this, we propose an effective hybrid deep learning method that combines Convolutional Neural Networks (CNNs) with Long Short-Term Memory (LSTM) networks and domain-driven feature engineering to incorporate occupancy and seasonality patterns. The dataset from an office building (research support facility) was used to perform a series of short-term (05-minutes), medium-term (daily), and long-term (weekly) forecasts. Experimental results show that the proposed forecasting approach demonstrates greater precision than the ten baselines including machine learning and deep learning methods in short, medium, and long time horizons. In particular, the model with feature engineering improved Root Mean Squared Error (RMSE) by 4.71% and Mean Absolute Percentage Error (MAPE) by 3.99% in the short term and saw even greater improvements for medium-term forecasts, with RMSE improved by 6.08% and MAPE by 5.37%, and for long-term forecasts, with RMSE improved by 2.58% and MAPE by 0.33%, respectively. The outcomes here offer significant benefits to the science community, supporting the development of such smart campuses with a specific, understandable methodology for energy forecasting. The methodology enables the application of data-driven energy management, cost optimization, and sustainability planning in educational spaces. The findings also provide valuable insights for deploying Artificial Intelligence (AI)-driven infrastructure management systems in academia and institutional energy governance.

Index Terms: Smart Campuses, Energy Forecasting, Energy Efficiency, Deep Learning, Educational Infrastructure, Sustainable Energy Management.

This work is open access and licensed under the Creative Commons CC BY 4.0 License.

1. Introduction

The need for an intelligent and sustainable power system has increased as global energy consumption continues to rise at an accelerating rate. The most effective way to bridge the energy gap between supply and demand is through the implementation of smart grids (SGs). Energy usage forecasting is a research area that has become an essential tool for SGs. According to the authors in [1,2,3], accurate energy forecasting is crucial for energy procurement, revenue analysis, capital spending, and market regulation. Some scientists were deterred by uncertainties in projections. New techniques, however, are always being invented to make more precise and trustworthy projections of energy consumption. Smart grids employ forecasted energy use data to plan their energy-producing assets, manage distributed energy management systems, initiate demand response systems, and adjust real-time energy market prices. As a result, smart grid operations become more stable, cost-effective, and flexible as energy forecasting improves.

In addition, residential, commercial, and industrial buildings use 20–40% of the world's energy, according to [4]. The exact measurement of energy used for certain tasks like lighting, heating, and cooling, as well as other crucial operations in residential, commercial, and institutional buildings, is included in the concept of energy consumption in buildings. Buildings in China and India consume 37% [5] and 35% [6] of the world's energy, respectively, making them a key focus for sustainability and energy efficiency initiatives [7]. Therefore, energy usage forecasting is again crucial for a smart building energy management system.

Due to the aforementioned reasons, our study focuses on improving energy efficiency in office buildings, since, considering the number of hours they are in use, they are among the most energy-intensive spaces. Reducing energy usage using algorithms based on artificial intelligence approaches is one of the goals of our study area for the general public in the years to come. In the literature, several studies have been conducted for commercial buildings' energy forecasting. There are three types of forecasting models used: machine learning, deep learning, and hybrid models. Forecasting energy usage is a time series problem. According to [8], standard machine learning approaches are unable to effectively anticipate energy consumption because of the seasonal volatility and irregular trends in time series data. Therefore, traditional machine learning (ML) is flawed for complex systems. Nonetheless, deep learning (DL) hybrid techniques have been extensively researched for energy forecasting and have demonstrated increased accuracy [9].

Recently, there has been a lot of interest in two types of DL neural networks: convolutional neural networks (CNNs) and long short-term memory (LSTM) networks. The authors in [10] looked into how well the CNN model predicted a residential building's electricity usage. The outcomes of the CNNs were contrasted with those of artificial neural networks (ANNs), support vector machines (SVM), and LSTM. The CNN model performs better than SVM and is on par with ANNs and LSTM. The goal of [11] was to forecast daily electricity use for a month in advance by taking into account four distinct household daily datasets. Their experiment showed that, in terms of average prediction accuracy and efficiency, the ConvLSTM model performs better than LSTM.

The authors in [12] suggested a hybrid model based on artificial neural networks for estimating power generation and day-ahead load. Ten months of data from commercial and industrial sites, gathered between July 2011 and April 2012, were used to evaluate the model. To solve the issue of short-term energy consumption forecasting, the authors in [13] presented a stacking multi-learning ensemble technique in a different study. The concept was evaluated using the 52-year Global Oil Consumption dataset (1965–2016). The authors in [14] suggested a hybrid model using SVR and Chaotic Sequence and Simulated Annealing to anticipate monthly electric load data from northeast China and New York City. Both sets of data span the period from January 2004 to October 2005. For electric load forecasting, the authors in [15] presented a novel hybrid approach that incorporates SVR, auto regression, and the differential empirical mode decomposition (DEMD) strategy. Two datasets were used to test the resulting approach. The first one includes New South Wales' half-hourly electric load data, and the second one uses the New York Independent System Operator's hourly electric load data.

Despite improvements in hybrid deep learning models, some crucial research gaps exist: most existing approaches failed to capture long-term dependencies and local patterns simultaneously; capturing both could further enhance the accuracy of the model. Most of the hybrid deep learning models were evaluated at a single time granularity. The effectiveness of these models was not tested in different time horizons – short-term, medium-term, and long-term. Also, most of the research work in this area did not consider the external variables that may affect the performance of the model. For example, occupancy patterns can affect the energy usage in the office building because of working hours, holidays, and weekends.

The problem statement is: to design an efficient energy forecasting model for an office environment that could include these factors and behave the same in all time granularities. Therefore, in this study, we present a hybrid deep learning model based on CNN and LSTM for energy forecasting using the office building dataset. Although the dataset originates from the National Renewable Energy Laboratory's (NREL's) Research Support Facility (RSF), it closely resembles the operating features of research labs in smart campuses. Modern university campuses have research labs and office spaces with a high density of plug loads, specialized equipment, continuous background power usage, and irregular occupancy patterns. These conditions are more similar to the RSF environment than to traditional classrooms or dorms. Therefore, the RSF dataset is a realistic substitute for assessing energy consumption forecasting models meant for smart campus research facilities. In such facilities, detailed device-level measurements are rarely available

but are very valuable for energy optimization. While CNN–LSTM architectures have been widely used in general time-series forecasting, their application to device-level energy use in research-focused smart campus environments remains limited. The novelty of this work lies not only in proposing a new architecture, but also in the adaptation and methodological design (like multi-horizon evaluation and engineered temporal features used as occupancy proxies for improved input to the deep hybrid architecture) tailored to a specific research laboratory facility within a smart campus context**.** To the best of our knowledge, this is rarely used for energy forecasting in educational buildings. We found that the hybrid model with engineered features outperformed all the baseline deep learning models across all time horizons.

Thus, the main contributions of our research work are:

- A CNN-LSTM hybrid deep learning model is proposed to analyze plug-level consumption data from a research laboratory office facility that operates with standard building occupancy patterns and equipment usage both during scheduled times and through different occupancy rates.
- We design engineered calendar-based time-context features, which include hour, weekday/weekend, and night indicators, and month, to track schedule-driven activity patterns because direct occupancy sensing is not available. These features act as occupancy proxies for human activity cycles.
- We conduct forecasting for short-term, daily, and weekly time periods by using resampled 5-minute device-level data to evaluate model performance at different time intervals.
- We assess multiple metrics, which include root mean squared error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and coefficient of determination (R^2), through 5-fold cross-validation to study model performance across various time periods.
- We use paired t-test and Diebold–Mariano (DM) testing methods to evaluate the statistical significance of our model improvements with respect to engineered and raw features; the improvements show consistent results despite their minor numerical increase.
- The proposed model undergoes testing against all baseline models to assess its performance.

The combined research improves comprehension of deep learning energy forecasting methods applicable to research-focused smart campus buildings and provides a reusable system that enables detailed energy consumption forecasts and energy control solutions.

1.1. Outline

The rest of this paper is organized as follows: Section 2 introduces a variety of innovative work done in the form of related work. Section 3 describes the mathematical expressions of deep learning models used to design the suggested hybrid model. It also describes the overall structure of the proposed methodology used along with the data description and pre-processing methods, hardware and software configurations, hyperparameter details, and performance indicators. Section 4 presents the results and discussion of the suggested model with the impact of external variables on the performance of the model in all time granularities. The conclusion and future scope of the paper are presented in Section 5.

2. Literature Review

Energy usage prediction methods are classified into several groups based on the end energy uses. This classification includes primary energy sources, such as natural gas, electricity, steam load usage, air conditioning, heating, and ventilation [16]. The authors in [17] distinguished two main classes based on the model's application. The first one is about building energy optimization and demand response management, and the second one is regarding the design of building parameters, which includes energy planning with its impact on the environment.

A variety of data-driven methods, from contemporary deep learning architectures to traditional statistical regression, have shown promise in estimating building energy usage. Four models—specifically, autoregressive moving average models with exogenous inputs, or ARMAX—were assessed by [18]. The authors in [19] suggested a system that makes use of a metaheuristic-based optimized least squares support vector regression model that uses the firefly algorithm for optimization (MetaFA-LSSVR) and a seasonal autoregressive integrated moving average model (SARIMA). The model provided an extremely precise and reliable day-ahead projection of building energy demand with an overall error rate of 1.18%.

Kim et al. [20] presented a comparative study on short-term electricity load forecasting for institutional buildings. The data was hourly consumption data collected from a university campus in Seoul. The authors examined a range of statistical approaches, including ARIMA, ARIMA-Generalized Autoregressive Conditional Heteroskedasticity (ARIMA-GARCH), and regression-ARIMA, alongside an ANN with exogenous inputs. To improve predictive performance, exogenous variables such as temperature (via degree-day indices), holiday schedules, and weekend indicators were included in the feature engineering step. They demonstrated that the ANN model, specifically the variant non-linear autoregressive with exogenous input (NARX), outperformed statistical models in all forecasting horizons from one hour to one day ahead. The paper also pointed out the pragmatic benefit of using a moving-window

simulation approach, which is closer to actual forecasting scenarios. The work is limited to short-term horizons, and there is also no scope for external variables to improve the accuracy.

Pinto et al. [21] proposed an office-level short-term electricity consumption forecasting using hybrid learning methods, namely Random Forests, Gradient Boosted Regression Trees, and an adapted Adaboost.R2. The data was taken from the Instituto Superior de Engenharia do Porto/ Research Group on Intelligent Engineering and Computing for Advanced Innovation and Development (ISEP/GECAD) campus in Portugal. The authors also employed lagged features and environmental variables such as temperature, humidity, and luminosity. The hybrid model outperformed other ensembles and classical models (including SVMs and fuzzy systems). While this study offered solid evidence that tree-based ensemble methods can deliver high performance in office building load forecasting, the work did not exploit deep neural architectures capable of extracting hierarchical temporal features.

Chitalia et al. [22] proposed a robust short-term electrical load forecasting framework for commercial buildings by integrating deep recurrent neural networks with unsupervised clustering and evaluated it across five building types in multiple geographic regions (Bangkok, Hyderabad, Virginia, New York, and Massachusetts). The approach investigated nine hybrid recurrent neural network (RNN) variants (including LSTM, BiLSTM, and attention-augmented architectures) paired with k-means clustering. It was done to partition operational regimes, which improves predictive performance under diverse building conditions. The authors further examined the influence of training data length, temporal resolution (e.g., 15-minute vs hourly), and weather forecast uncertainty on forecasting accuracy, finding that acceptable performance may be achieved with as little as one month of data and that higher data resolution can provide ~30% gains in hour-ahead forecasts. Their hybrid models delivered 20%–45% improvement over prior benchmarks for both hour-ahead and 24-hour forecasts, indicating strong robustness across building types and locations.

Pallonetto et al. [23] examined electricity demand forecasting in commercial buildings with an eye toward enabling demand response programs. Their work compares the LSTM and SVM models that were trained on the same actual building load data. The authors found that LSTM performed better when trained on sufficient data, while SVM is the better choice when there is limited training data or when computation cost is the primary consideration. Their findings also emphasized the improvement in forecasting performance obtainable through pre-processing and the judicious choice of exogenous features. However, this work was limited to relatively standard models and did not explore more powerful hybrid deep learning architectures.

Hwang et al. [24] proposed a hybrid two-step forecasting framework for commercial buildings that tackles the challenge of incomplete system data by inferring hidden features like heating/cooling equipment types via DTW-based clustering and by introducing a novel compensation parameter (Q-value) combining building area and temperature variation. The authors obtained the data from twenty-eight commercial buildings in Gwangju, South Korea, and compared models including SARIMAX, SVR, ANN, deep neural networks (DNN), and LSTM. The findings revealed that the augmented feature models outperformed baseline configurations. This work was limited to daily and monthly based prediction horizons only, and the authors did not explore advanced hybrid deep learning architectures.

Wen et al. [25] developed a gated recurrent unit-temporal convolutional network (GRU–TCN)–Attention hybrid model, which delivers better accuracy in short-term load forecasting than existing baseline methods across multiple load testing datasets. The model demonstrates better performance, yet its capacity to generalize results is restricted because testing used only a small selection of datasets, and the deep learning system design required multiple components, which made the model harder to implement and understand compared to basic systems.

Li et al. [26] introduced a hybrid model that combines a residual and an attentive LSTM with a TCN for short-term load forecasting in integrated energy systems. The model attained better results with the MAPE score of 2.35% on a campus-sized dataset because it successfully captured both short-term fluctuations and extended temporal patterns. The research has two existing gaps: the model's performance needs to be tested in new locations, and its computational requirements exceed current research limits.

Bai et al. [27] developed a hybrid TCN–BiLSTM model for industrial equipment fault prediction, which shows that the combined use of TCN multi-scale temporal feature extraction and BiLSTM long-sequence modeling techniques enhances predictive accuracy on the AI4I maintenance dataset, although it requires extensive computational resources because of its intricate model design.

Parizad et al. [28] created a particle swarm optimization **(**PSO)-optimized eXtreme Gradient Boosting (XGBoost) hybrid model to forecast household energy demand and electricity prices together. This model showed much better accuracy compared to traditional ML baselines. However, its dependence on a single-year dataset and lack of deep temporal modeling restrict its generalizability.

Bai [29] presented a genetic algorithm (GA)-optimized attention-based LSTM model that predicts residential energy consumption with reduced error rates in the short-term horizon compared to all traditional data-driven methods. The model uses a single-household dataset, which creates some issues that prevent it from being used in other situations, and its operational costs make it infeasible to deploy in real life.

Özüpak and Mansurov [30] proposed an RNN-LSTM hybrid model to predict electricity demand for short and medium horizons. The system outperforms standard deep learning models. The research needs to continue because two issues exist: the system requires more processing power, and its performance depends on specific regional data.

Mariano-Hernández et al. [31] examined forecasting of hourly electricity consumption in buildings under small data conditions. Using data from two university buildings with four years of historical records (15-minute resolution), they compared shallow models (Random Forest (RF), XGBoost) and deep models (CNN, TCN) under two forecasting strategies: (a) forecasting all hours of the next 24 h at once; (b) building separate models for each hour of the day. Their experiments showed that for horizons of one week, only about six months of past data suffices to reach MAPE below ~7-10%, increasing to about ten months for one-month-ahead forecasts. The authors also found that when consumption behavior is stable, shallow models often outperform deep models for small datasets, but with sudden changes in load patterns (due to efficiency upgrades, renewable integration, etc.), deep learning models adapt better. While this work offered valuable insights into data size requirements and algorithm selection under limited history, it does not explore hybrid architectures. The advanced architectures of Transformers and Neural Basis Expansion Analysis for Time Series (N-BEATS) have achieved the best performance results in their field, but their high computational demands and extensive data requirements create limitations on their use in resource-limited smart-building settings. The current need for real-time systems exists because lightweight hybrid models provide better solutions for deployment purposes. Some of the recent research in the area of smart commercial buildings' energy usage forecasting using hybrid methods is analyzed and provided in Table 1.

Table 1. Summary of some recent studies for energy forecasting in commercial buildings using deep learning & hybrid models.

Ref.	Models Used	Dataset Used	Forecast Horizon	Contributions	Advantages/Key Results	Research Gaps
[33]	LSTM and GRU recurrent neural networks	Data were provided by the energy management company Emececuadrado (Badajoz, Spain). The paper uses hourly active energy consumption (1 Sep 2016 – 1 Jul 2021; ≈40k points) and, for the multivariate case, reactive energy, outside temperature, and outdoor humidity. (No public data link is available)	Short-Term	• The authors proposed a simple, practical hybrid pipeline combining signal decomposition with RNN forecasting (LSTM & GRU) for multi-step hourly forecasts in a hospital • Provided a detailed experimental setup (sliding window training, metric analysis) and reports best-case performance (MAPE & RMSE)	• Best reported: LSTM + preprocessing (multivariate) achieved MAPE 3.51% and RMSE 55.06 (units: kWh • Both LSTM and GRU performed similarly overall; preprocessing gave largest gains in the multivariate scenario.	• The empirical mode decomposition (EMD) family is heuristic; paper did not deeply analyze sensitivity to decomposition parameters. • No comparison was done with hybrid architectures. • No dataset link was provided. • The model was not able to capture the local contextual information.
[34]	Seq2Seq RNNs	The data was collected from four commercial office buildings in the United States. The period of data and exact public link are *not* provided in the paper.	Short-Term	• The authors proposed a simple architecture based on seq2seq RNN for energy forecasting and provided a guidance on the range of hyperparameters used for training. • The authors also provided some important insights about how context length impacts performance of the model.	• Provided an empirical evidence for minimum data requirements (both in terms of duration and context length) useful for practitioners	• Dataset not public: makes reproducibility / benchmarking by others harder. • Only some standard exogenous factors (outdoor temperature) were used. Other exogenous factors (occupancy, internal loads, solar radiation, humidity, schedule etc.) were not deeply exploited. • The model was evaluated in short-term horizon only.
[35]	RF-LSTM	The authors utilized hourly energy consumption data from five different buildings, but they have not provided a direct link to this dataset	Short-Term	• The authors introduced a novel hybrid model combining RF and LSTM, enhanced by complementary ensemble empirical mode decomposition (CEEMDAN) decomposition, to improve building energy consumption forecasting accuracy.	• The hybrid RF-LSTM model outperformed traditional methods in forecasting accuracy.	• The model did not incorporate external factors such as weather conditions, occupancy patterns, or HVAC system data, which could influence energy consumption. • The decomposition and hybrid modeling process may be computationally intensive, potentially limiting real-time application. • The model was not

						able to capture the local patterns.
[36]	RF and Bi-LSTM	The data was collected from a civil public building in Dalian City, China. The dataset includes various factors influencing energy consumption, such as temperature, humidity, and occupancy levels. The dataset is not publicly available.	Short-Term	• The authors proposed a novel hybrid model that uses Entropy-Weighted K-means, and RF for features selection, BiLSTM for forecasting, and sparrow search algorithm (SSA) for model optimization to enhance the accuracy of building energy consumption predictions	• The proposed hybrid model outperformed all other baseline models in terms of accuracy and robustness.	• The model did not incorporate external factors such as energy prices or policy changes, which could influence energy consumption patterns. • The model was evaluated in short-term horizon only and not able capture the local patterns.
[37]	LSTM - double attention	The data for this study was recorded at a green building located in Shenzhen, China.	Short-Term	• The authors proposed a novel hybrid model that combines the double attention mechanisms with LSTM and evolutionary algorithms to enhance the accuracy of building energy consumption predictions.	• The authors used the attention mechanism that allows the model to focus on the most relevant features, enhancing prediction performance. • The hybrid model exhibits strong robustness across different seasons, maintaining consistent performance	• The model did not incorporate external factors and was evaluated in short-term horizon only. • The complexity of the hybrid model may reduce its interpretability, making it challenging for practitioners to understand the decision-making process. • The model was not able capture the local patterns.

3. Methodology

This section describes the proposed methodology for energy forecasting in smart office buildings using hybrid deep learning models. The hybrid model combines the advantages of two models: CNNs and LSTM networks. The model is tested and trained using 5-fold cross-validation, with test metrics comprising RMSE, MAE, MAPE, and R^2. The overall structure of the proposed methodology is given in Fig. 1. In the proposed framework, the energy usage data from plug-in devices is collected and preprocessed in the first step. After that, the hybrid model is constructed using the two components- CNN & LSTM. CNN consists of one layer of convolution (1D CNN) with kernel size=3 and number of filters=64, and one layer of max pooling with pool size=2. The LSTM component consists of two layers of LSTM, with 128 and 64 units, respectively. It also uses dropout layers in between these LSTM layers with a dropout rate of 0.25. To train this model, the k-fold cross-validation re-sampling process is used to create the train and test datasets, in which 80% of the data is used for training and 20% of the data is used for testing. The grid search method is used to find the ideal parameters for the number of epochs, batch size, dropout rate, lag columns, and the count of LSTM units. Finally, the model is evaluated using various performance indicators. The detailed description of the proposed methodology is provided below.

3.1. Data Collection

The data is crucial to train any deep learning model. This dataset was collected from the office building, and it was used to assess our hybrid deep learning based energy forecasting model.

The dataset used in this work was obtained from the online public repository: Mendeley Data [38], which contains the energy observations from devices used in the Research Support Facility (RSF—a zero-energy office building) at the National Renewable Energy Laboratory (NREL). The measurements were recorded over around 3 months, from October 01, 2017, to December 31, 2017, at 05-minute intervals. It has 26496 instances and 16 features with one target feature, i.e., Skyspark, which represents the power consumption data from the plug load submeter in RSF.

The educational building considered here functions as a contemporary university campus facility that combines academic functions with educational spaces, research areas, and laboratory environments. University buildings frequently host research labs such as the RSF and NREL to support engineering and science programs, and such spaces are formally recognized within Department of Energy (DOE)/American Society of Heating, Refrigerating, and Air Conditioning (ASHRAE) building classifications as standard components of higher education infrastructure. The energy consumption patterns of these laboratories exceed those of regular classrooms because their specialized electrical and computing equipment requires more power than standard classroom equipment, which operates with known power requirements. The educational building features both RSF-type research laboratories and NERL-type research laboratories, which fulfill architectural design requirements that universities need for their energy modeling research

according to current standards. This justifies the use of this data as a representative dataset for the smart-campus energy forecasting context. The in-depth explanation about the features and data features statistics is provided in Table 2 and Table 3, respectively.

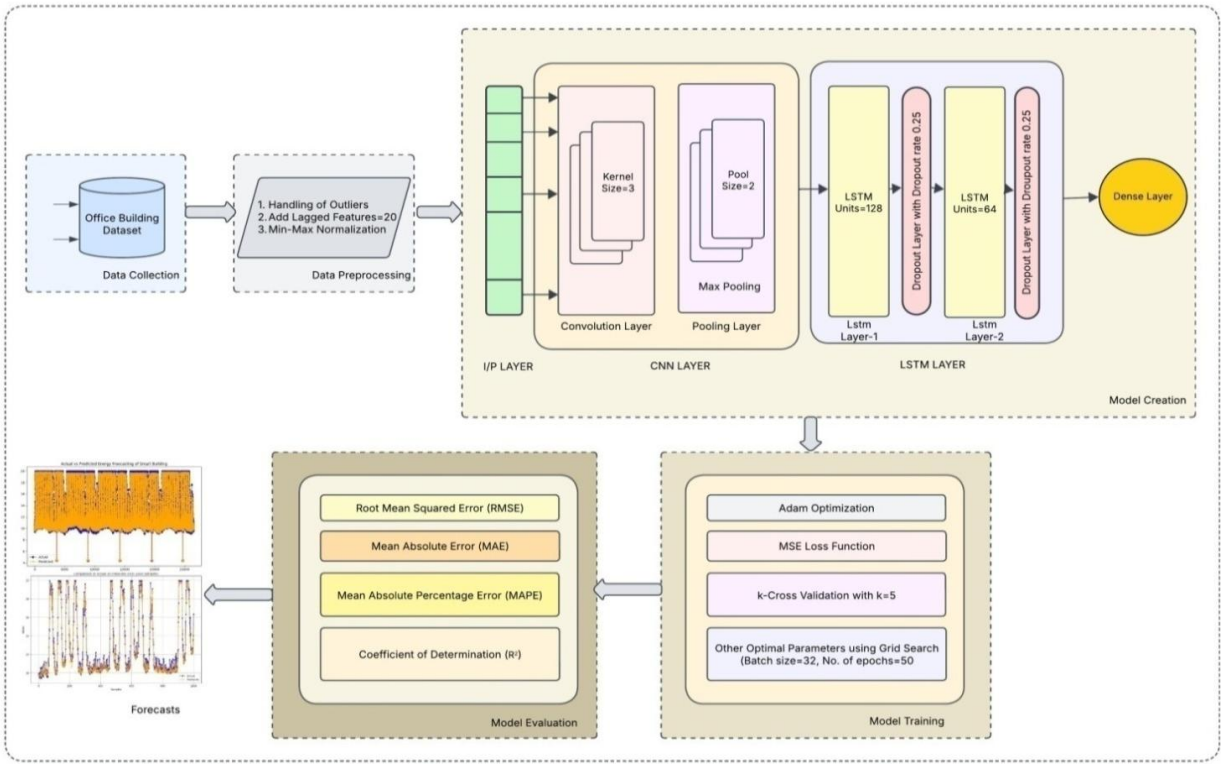


Fig. 1. Overall structure of the proposed methodology.

Table 2. Features description of office building dataset.

S.No.	Attributes	Measures	Depiction
1.	datetime	dd-mm-yyyy hh:mm	Date and time at which the energy measurements were taken.
2.	**Skyspark**	kW	The total aggregated plug load energy consumption in office building.
3-4.	Coffe.Maker & Microwave	kW	The power consumed by the kitchen appliance in the office building.
5-6.	Copier & Printer	kW	The power consumed by the copier & printer used in the office building.
7-8.	Desktop & Laptop	kW	The power consumed by the computer systems including desktops and laptops.
9-11.	Monitor & Display screens-Monitor, TV and Projector	kW	The power consumed by the monitors, TV, and projectors used in the office building.
12.	Water.Boiler	kW	The power consumed by the water boiler used in the office building.
13-16.	Plug-in Devices- Lamp, Headset, Phone.Charger, Video. Conference.Camera	kW	The power consumed by the plug-in devices like lamps, headsets, phone charger, and video conference camera.

Table 3. Descriptive statistics of data features.

Features	Count	Mean (kW)	Std. Deviation (kW)	Min (kW)	Max (kW)
Skyspark	26496	12.5674	3.913659	0	26.624
Coffe.Maker	26496	0.006658762	0.051815226	0	0.844113333
Microwave	26496	0.043585	0.167671	0.00472	2.8216
Copier	26496	0.039166	0.053127	0	0.91542
Printer	26496	0.027893	0.038726	0	0.86228
Desktop	26496	0.134979	0.029609	0	0.410367
Laptop	26496	0.102443	0.1135	0.002113	0.640474
Monitor	26496	0.101156	0.145159	0.005721	0.667159
TV	26496	0.049709	0.050026	0	0.458376
Projector	26496	0.039215	0.10131	0	0.684897
Water.Boiler	26496	0.03937115	0.17572851	0.00033333	2.56971667
Lamp	26496	0.013781	0.003118	0.006743	0.047113
Headset	26496	0.005684	0.000208	0.001753	0.007495
Phone.Charger	26496	0.0002	0.0005	0	0.0067
Video.Conference.Camera	26496	0.208168896	0.168942233	0	0.370599608

The count of all the features is the same; therefore, we can assume that there are no missing values in the data. The high value of Skyspark's standard deviation in the above table represents the high variability in the aggregate energy usage. Fig. 2 shows the distribution of all the features, and it can be observed that the distributions are irregular and indicate the presence of outliers in the data.

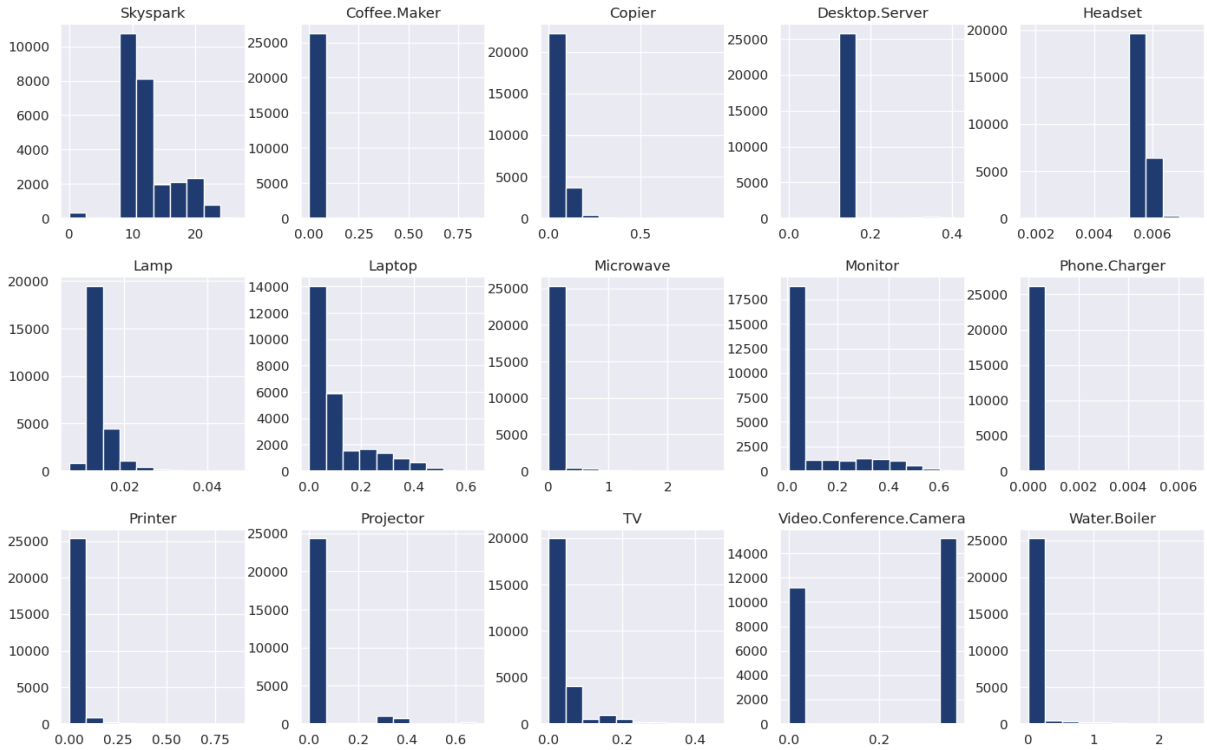


Fig. 2. Feature distributions.

3.2. Data Preprocessing

This section describes the data preprocessing methods used to clean the data. The sequence of methods used for preprocessing is as follows.

3.2.1. Handling of Outliers

In this first step of preprocessing, the outliers were addressed using the Interquartile Range (IQR) method. In this approach, the IQR is calculated as the difference between the 75th percentile (QT3) and the 25th percentile (QT1), and the lower bound value (l) and the upper bound value (u) are derived accordingly. The mathematical expressions to calculate these values are illustrated using (1) and (2), respectively.

$$l = QT1 - (1.5 * IQR) \tag{1}$$

$$u = QT3 + (1.5 * IQR) \tag{2}$$

As in (3), any value (v) above this u and below this l is defined as an outlier.

$$v < l \; or \; v > u \tag{3}$$

This is the method to cap the outliers [39]. We did not remove the outliers in order to prevent the loss of useful information from the data.

3.2.2. Add Lag Features

The lag features were incorporated in the feature set to help the model learn short-term dependencies effectively. This was done using the grid search method with the test combinations of 20, 30, and 40. Out of these tested combinations, the best value was obtained with the number 20. Therefore, the number of lagged features (L_f) added in the given dataset is 20.

3.2.3. Min-Max Normalization

For normalization purposes, we used the min-max scaler that ensures equal feature contribution [40]. It normally scales the input features to the range between 0 and 1 as per the mathematical expression given in Eq. 4.

$$v'^{k}_{f} = \frac{v^{k}_{f} - v_{f,min}}{v_{f,max} - v_{f,min}} \tag{4}$$

where v'^{k}_{f} and v^{k}_{f} represent the scaled and ground truth values for feature 'f' at time instant 'k', respectively. $x_{f,max}$ and $x_{f,min}$ are the maximum and minimum ground truth values of feature 'f', respectively.

3.2.4. Temporal Aggregation for Medium-Term and Long-Term Forecasting

The load dataset, which was originally gathered from the National Renewable Energy Laboratory Research Support Facility (RSF), provides 5-minute interval measurements. The proposed CNN-LSTM model evaluation needs raw data to be converted into lower time intervals because the model needs to be tested across multiple forecast periods. The preprocessing workflow included temporal aggregation as a necessary step. To enable medium-term forecasting, we resampled the 5-minute load data into daily aggregated values using temporal averaging. This step reduces high-frequency variability, captures daily load characteristics, and creates a smoother series suitable for multi-step prediction each day. For long-term forecasting, we further resampled the data into weekly averaged values, again using mean-based temporal resampling. Weekly aggregation highlights broader consumption patterns and allows us to improve the model's ability to forecast over longer periods within the 3-month dataset.

This aggregation strategy is consistent with established building energy forecasting practices, allowing the construction of medium- and long-horizon forecasting tasks without requiring a full year of data. It also ensures that the proposed model is evaluated on progressively lower-frequency signals while avoiding claims related to seasonal long-term forecasting, which cannot be assessed with the available dataset duration.

3.3. Model Creation

This section describes the deep learning models used to create our proposed hybrid model architecture. The proposed model uses the strength of two deep learning models, i.e., CNN and LSTM, to capture both local temporal patterns and long-range dependencies in building energy consumption data. The mathematical description of the proposed hybrid model is provided below.

Let the input multivariate time series be given in (5).

$$F = \{f_t \in \mathbb{R}^d \mid t = 1,2,\dots\dots,T\} \tag{5}$$

where f_t is the input feature vector at time step t (including lagged features), d is the number of features, and T is the sequence length. The following are the components of the proposed model.

- **CNN Layer**: It captures the short-term local patterns effectively. CNN assists in reducing the random noise while preserving the salient features in the energy data [41]. This is done by aggregating information through a 1D convolution layer and a 1D max pooling layer of the CNN architecture. The 1D convolution extracts local temporal patterns from the input sequence with the help of the mathematical expression given in (6).

$$h^{(c)}_{t,k} = \sigma\left(\sum_{n=1}^{ks} W^{(c)}_{k,n} \cdot f_{t-n+1} + b^{(c)}_{k}\right) \tag{6}$$

where ks is the filter (kernel) size, $W^{(c)}_{k.n}$ and $b^{(c)}_{k}$ are learnable weights and bias of the k -th filter, σ(·) is the ReLU activation function. This produces feature maps with K filters using the mathematical expression given in (7).

$$H^{(c)} = \{h^{(c)}_{t,k} \mid t = 1,2,\dots\dots,T, k = 1,\dots\dots.K\} \tag{7}$$

- **Max-Pooling Layer:** This layer is used to reduce dimensionality and focus on dominant local features. The mathematical expression for this layer is given in Eq. 8.

$$h^{(p)}_{t,k} = \max_{l \in \mathrm{P}} h^{(c)}_{t+l,k} \tag{8}$$

- **LSTM Layer:** The pooled CNN features are fed into an LSTM to capture the long-term dependencies effectively [42]. LSTMs can capture non-linear temporal dynamics such as seasonality, peak loads, and gradual consumption

shifts, which traditional machine learning models fail to learn. In our proposed model, 2 layers of LSTMs act as the memory backbone of the hybrid model, enabling it to learn complex, long-range temporal dependencies in energy data, which a CNN layer alone cannot capture. This is possible only due to the three gates used by LSTM. The mathematical description of these gates is given in (9-14).

$$FORG_t = \sigma\,(W_{FORG} \bullet [h_{t-1},\ x_t] + bias_{FORG}) \tag{9}$$

$$ING_t = \sigma\,(W_{ING} \bullet [h_{t-1},\ x_t] + bias_{ING}) \tag{10}$$

$$\widetilde{CDS_t} = tanh(W_{CDS} \bullet [h_{t-1},\ x_t] + bias_{CDS}) \tag{11}$$

$$OUTG_t = \sigma(W_{OUTG} \bullet [h_{t-1},\ x_t] + bias_{OUTG}) \tag{12}$$

$$CS_t = FORG_t \bullet CS_{t-1} + ING_t \bullet \widetilde{CDS_t} \tag{13}$$

$$h_t = OUTG_t \bullet \tanh(CS_t) \tag{14}$$

where $FORG_t$, ING_t,$OUTG_t$, are forget, input, and output gates respectively; $\widetilde{CDS_t}$ is the state of candidate cell; CS_t is the cell state; $h_t$is the hidden state; x_t is the input at time t; W_{FORG}, W_{ING}, W_{OUTG}, W_{CDS} are the weight vectors of all the gates respectively; $bias_{FORG}$, $bias_{ING}$, $bias_{CDS}$, $bias_{OUTG}$ are the biases in all the corresponding gates; σ is the activation function; tanh is the hyperbolic tangent function; is the element-wise multiplication operator.

- ***Dropout Layers***

The dropout function is one of the most effective strategies to regularize and prevent overfitting in neural network structures [43]. For this purpose, we used two dropout layers after each LSTM layer.

- ***Dense Layer & Forecast Output***

The final hidden state h_T is fed into the fully connected layer (dense layer). The mathematical description of this layer is given in (15).

$$y'_{t+1} = W_d.\ h_T + bias_d \tag{15}$$

where $y'_{t+1}$is the predicted value of energy consumption at the next time step, $W_d, bias_d$ are the weights and bias of the dense layer.

- ***Loss Function***

The model is trained by minimizing mean squared error (MSE). The mathematical expression for the same is given in (16).

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - y'_i)^2 \tag{16}$$

where y_i is the ground truth value of energy consumption, and y'_i is the anticipated value of energy consumption.

3.4. Model Training & Cross Validation Strategy

This section provides an in-depth explanation of the training process used for the proposed hybrid model. The hyperparameter tuning was done using the grid search method to select the best parameters. The data was partitioned into 20% of test and 80% of train data using the k-fold cross-validation (KF) (k=5) method, because it produces superior model generalization and prevents excessive model bias and variance, as demonstrated empirically in the literature [44]. This research uses a sliding-window approach to create its forecasting system which uses fixed past load values as input to predict future outcomes. The transition of time series data into lag-window samples creates a supervised learning matrix which eliminates its original sequential stream structure. The process of shuffling instances does not leak any future information because the windows maintain their original time-based arrangement. The current method follows the pattern established by recent deep-learning energy forecasting studies which use randomized KF to produce consistent and neutral performance assessment results [45,46]. The hyperparameters tuned for model training are listed in Table 4.

Table 4. Hyperparameters details

S.No.	Hyperparameters	Tested Possible Combinations	Best Values Selected
1.	No. of Lagged features	{20, 30, 40}	20
2.	LSTM	No. of layers: 1,2,3,4 No. of LSTM units: {32,64,128} Activation function: {"relu", "tanh"}	2 64, 128 (for both the layers respectively) tanh
3.	BiLSTM	No. of Bi-LSTM units: {32, 64, 128} No. of layers: 1 Activation function: {"relu", "tanh"}	64 1 tanh
4.	CNN	No, of filters: {32, 64} Size of kernel: {2, 3} No. of 1D-CNN layers: 1 No. of 1D Max pooling layers: 1 Activation function: {"relu", "tanh"}	64 3 1 1 relu
5.	XGBoost	No. of estimators: {300, 400, 500} Learning rate: {0.1, 0.15, 0.25} Max. depth: {2, 4, 6} Subsample: {0.6, 0.8, 1.0}	500 0.15 4 0.6
6.	NN	No. of layers: 2,3,4 No. of units per layer: {32,64,128} Activation function: {"relu", "tanh"}	2 64, 128 (for both the layers respectively) relu
7.	TCN	No, of filters: {32, 64} Size of kernel: {2, 3} Dilation rate: {1,2} No. of layers: 2 Padding=" causal"	64 3 1,2 2 " causal"
8.	Transformer	No. of attention heads: {2,4,6} Key vector dimensions: {16,32,64} FFN hidden size: {32,64,128} Learning rate= {0.01, 0.001, 0.0001}	4 32 64 0.001
9.	Batch size	{32,64}	32
10.	No. of epochs	{50, 100}	50
11.	Dropout rate	{0.15, 0.25. 0.35}	0.25
12.	Loss function	Mean Squared Error	Mse
13.	Optimizer	{Adagrad, SGD, Adam}	Adam

3.5. Model Evaluation

This section describes the performance indicators used to evaluate our proposed single-step energy forecasting model. The mathematical expression to compute the value of each indicator is provided in Table 5. These parameters assess our model's capability for generalization and forecasting results [47].

Table 5. Mathematical expression of performance indicators.

Performance Indicators	Description	Mathematical Expression
MAE	Mean Absolute Error	$MAE=\frac{1}{n}\sum_{i=1}^{n}\lvert v-v'_i\rvert$
RMSE	Root Mean Squared Error	$RMSE=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(v_i-v'_i)^2}$
MAPE	Mean Absolute Percentage Error	$MAPE=\frac{1}{n}\sum_{i=1}^{n}\left\lvert\frac{v_i-v'_i}{v_i}\right\rvert$
R^2	Coefficient of Determination	$R^2=1-\frac{\sum_{i=1}^{n}(v_i-v'_i)^2}{\sum_{i=1}^{n}(v_i-v''_i)^2}$
FT	Fitting Time- The time needed to fit or train the model on the given dataset	$T_f=T_{e_{t-fit}}-T_{s_{t-fit}}$
PT	Prediction Time- The time needed to predict the future data.	$T_p=T_{e_{t-pred}}-T_{s_{t-pred}}$

where n is data points number, v_i, $v'_i$and v''_i are the ground truth, predicted, and average value respectively.$\mathbf{T}_{s_{t-fit}}$, and $\mathbf{T}_{e_{t-fit}}$ are the start time and end time for training, respectively. $\mathbf{T}_{s_{t-pred}}$, and $\mathbf{T}_{e_{t-pred}}$ are the start time and ends for predictions respectively.

4. Results and Discussion

The experiment was conducted on highly configured computer systems equipped with CPU: Intel Core i7-12700 along with an extended NVIDIA Tesla P100 GPU (16 GB memory) and 29 GB RAM. The implementation was done in

the programming environment provided by Python 3.10.12 with deep learning packages like TensorFlow 2.17.1 and Keras 3.5.0. The proposed model was compared with ten baseline machine learning & deep learning models to test the effectiveness of the model. The baseline models considered for the comparison are: XGBoost, NN, LSTM only, CNN only, Stacked LSTM (2 and 3 layers), Bi-LSTM, Bi-LSTM+Stacked LSTM, TCN only and Transformer based model because these are the strong representatives of machine learning and deep learning methods. Table 6 and, Fig. 3 present the comparative performance of different deep learning architectures for short-term energy consumption forecasting after applying pre-processing. The evaluation metrics considered are RMSE, MAE, MAPE, and R^2.

Table 6. Short-term performance comparison of the proposed model (without engineered features-raw) across all the baselines.

Model Type	RMSE ↓	MAE ↓	MAPE (%) ↓	R^2 ↑
Pre-processing + XGBoost	0.5716	0.3914	3.07	0.9734
Preprocessing + NN	0.5632	0.4074	3.41	0.9741
Pre-processing + LSTM	0.5845	0.4076	3.15	0.9720
Pre-processing + Stacked LSTM (2 layers)	0.5788	0.4117	3.21	0.9726
Pre-processing + Stacked LSTM (3 layers)	0.5692	0.3929	3.07	0.9736
Pre-processing + Bi-directional LSTM	0.5866	0.4010	3.14	0.9719
Pre-processing + Bi-directional LSTM + Stacked LSTM (2 layers)	0.6983	0.5317	4.51	0.9292
Pre-processing + CNN	0.6799	0.4842	3.68	0.9622
Pre-processing + TCN	0.7612	0.4869	4.75	0.9483
Preprocessing + Transformer	3.5035	2.8346	22.92	-0.0001
Proposed Model (Pre-processing + CNN+LSTM)	**0.5610**	**0.3891**	**3.01**	**0.9743**

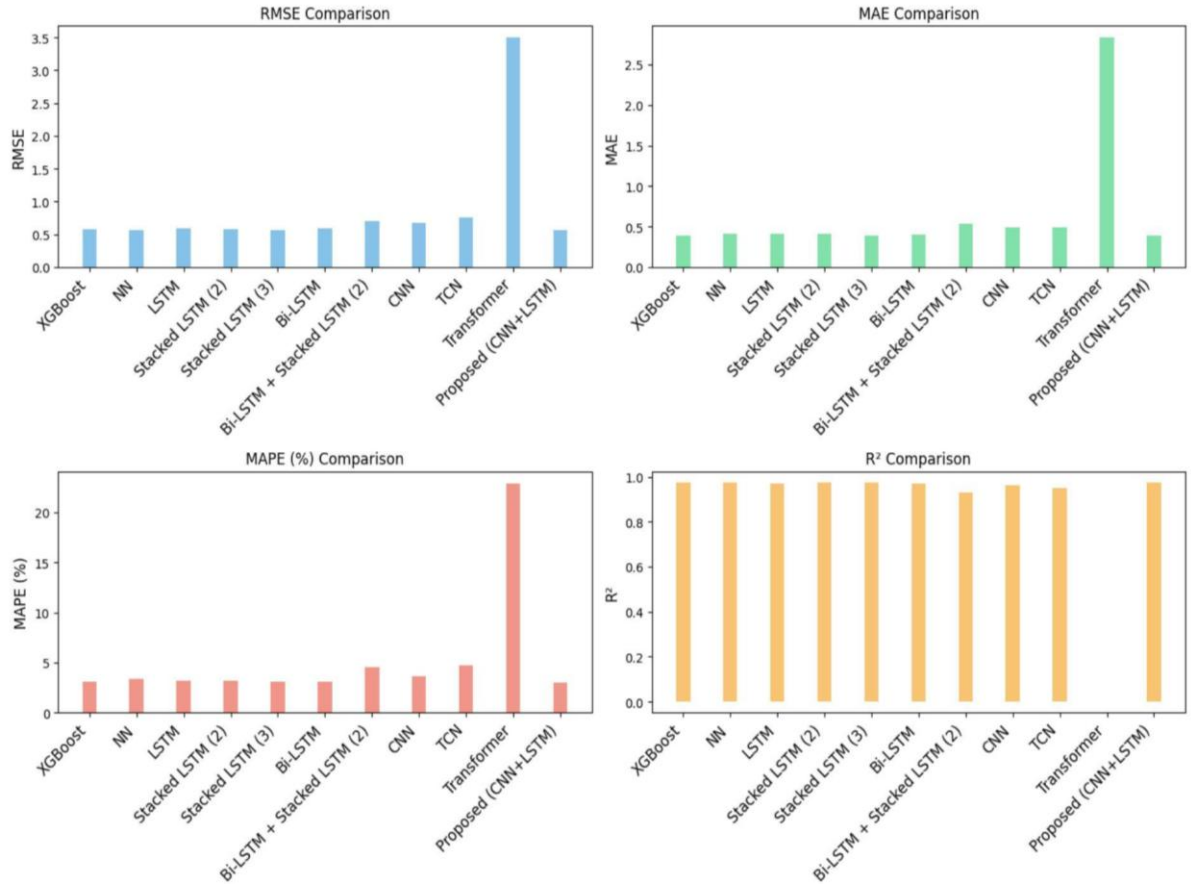


Fig. 3. Short-term comparative performance of the proposed model with all the baselines (RMSE, MAE, MAPE & R^2 comparison).

During testing the baseline simulations, it has become evident that XGBoost and Neural Network model (NN) attain good accuracy with RMSE of 0.5716 and 0.5632 respectively. This is because the size of the data is not large and these models usually perform well with small data size. The single-layer LSTM also attains good accuracy and generalization with RMSE of 0.5845, MAE of 0.4076, MAPE of 3.15%, and R^2 value of approximately 0.97. Yet another equally good prediction model built using Bi-LSTM achieves RMSE of 0.5866, MAE of 0.4010, and MAPE of 3.14%. More layers do add slight efficiency to forecasting, reducing RMSE to 0.5788 and 0.5692, while MAPE becomes 3.21% and 3.07% for stacked LSTM with 2 and 3 layers, respectively, thus pointing to further extraction of temporal information with layers. However, deep networks like Bi-LSTM combined with stacked LSTM result in a

decline in performance (RMSE ≈ 0.6983 and MAPE ≈ 4.51%), thus revealing that too much network depth increases overfitting possibilities and complexities. The single CNN model showed relatively poor performance (RMSE = 0.6799, MAE = 0.4842, and MAPE = 3.68%), denoting the inability to recognize long dependencies in sequential energy data. The TCN model is less accurate with the RMSE of 0.7612 because the TCN tends to overfit early if the data is not large enough and fails to capture small fluctuations. Similarly, Transformer based model baseline underperformed (RMSE ≈ 3.50, $R^2 \approx 0$), which is consistent with prior studies showing that Transformer architectures require much larger datasets and longer sequence lengths to stabilize training [32]. However, the proposed hybrid CNN+LSTM model yields the best performance record, being superior to others by all parameters; RMSE = 0.5610, MAE = 0.3891, MAPE = 3.01%, and the highest R^2=0.9743, showing that a combination of CNN's capability to learn local temporal patterns with the LSTM's capability to model sequential dependency enhances forecast accuracy. The hybrid architecture, therefore, balances local feature learning and long-term dependency modeling and outperforms individual CNN and LSTM models.

Overall, these results confirm that the CNN+LSTM hybrid model gives better predictive accuracy and resilience, affirming its efficacy in smart building energy forecasting solutions. Fig. 4 illustrates the similarity between the actual and forecast energy consumption values generated through the use of the proposed CNN+LSTM hybrid model for smart building forecasting. The predicted curve (orange) is very close to the actual consumption (blue), representing the high accuracy and reliability of the model. The close following across the whole-time horizon ensures the capability of the model to identify peak demands as well as normal consumption patterns effectively. Infrequent deviations are insignificant, demonstrating the good generalization ability of the model in managing variable load patterns. These findings indicate the robustness of the given model in actual energy forecasting practices for effective decision-making and sustainable energy operations.

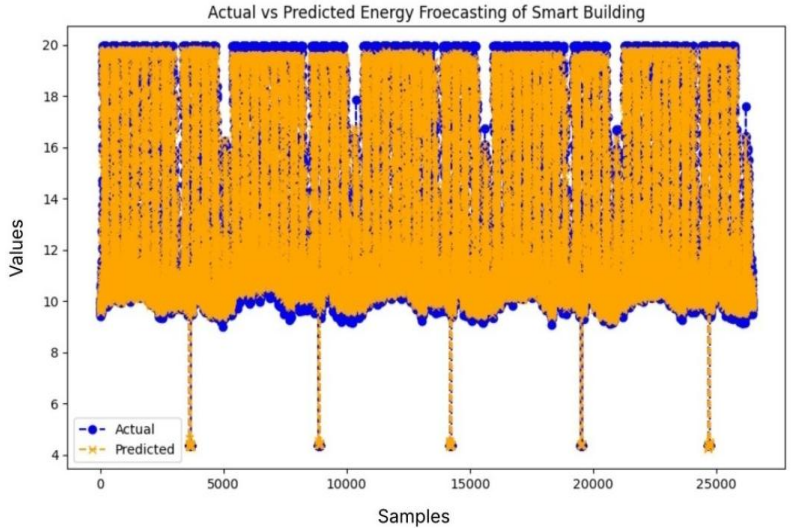


Fig. 4. Actual vs. Predicted energy load for the proposed model.

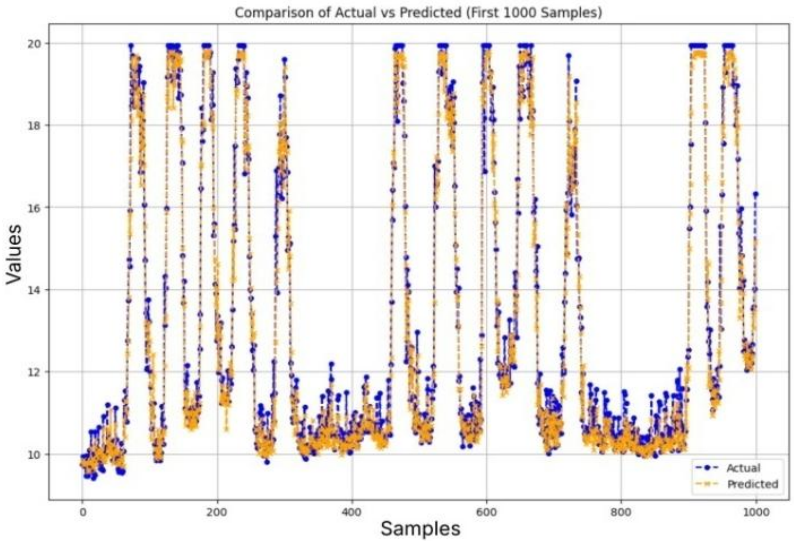


Fig. 5. Actual vs. Predicted energy load (First 1000 samples) for the proposed model.

Fig. 5 and, Fig. 6 give a magnified view of the first 1000 samples and the last 1000 samples, respectively, showing graphs of the actual and predicted energy consumption. The predicted series (orange) is aligned with the actual consumption values (blue) and captures sharp local peak values, sudden changes, and valleys of low consumption. This indicates that the model proposed can adapt to short-term variability and sudden demand spikes, while remaining consistent when the region is stable. This close overlap between curves confirms the robustness of the model in depicting both types of load patterns, i.e., dynamic and steady-state, which are required for short-term forecasting.

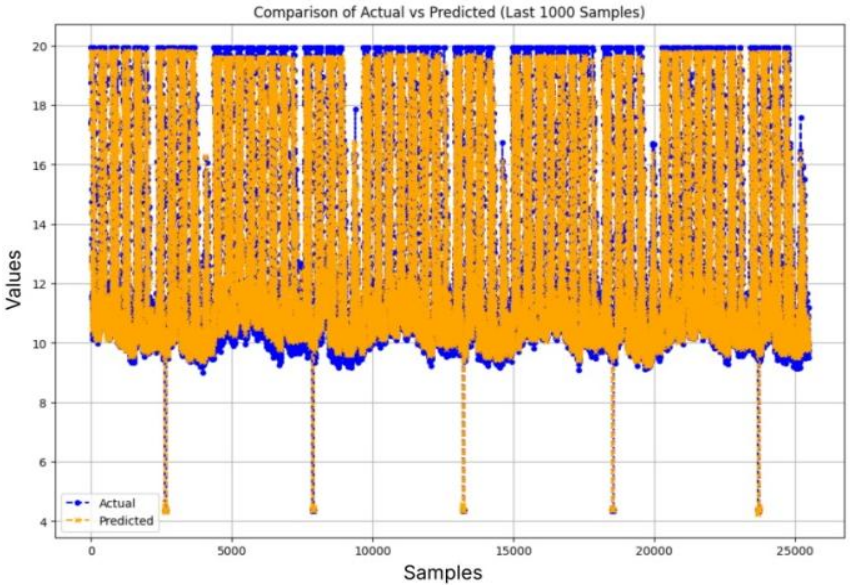


Fig. 6. Actual vs. Predicted energy load (Last 1000 samples) for the proposed model.

The plot in Fig. 7 shows the training loss curve of the proposed model across 50 epochs. At the initial stage of training, one expects a relatively high loss (around 16) since the model has been initialized with random weights. The loss decreases sharply within a few epochs, whereby one can intuitively say the model is learning some meaningful data patterns. After the steep decline, the loss decreases slowly and converges smoothly. After the 40th epoch, the loss remains below one, showing that the model has sufficiently reduced the error and reached a steady state for learning, thus not overfitting.

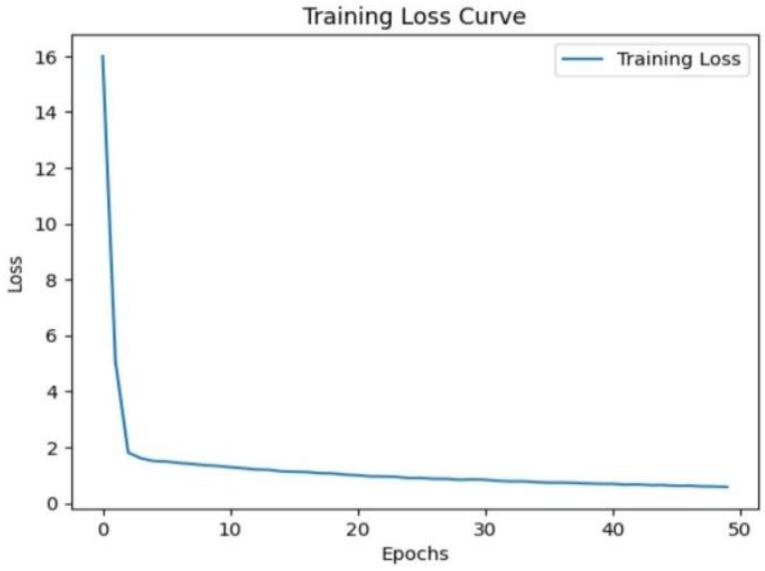


Fig. 7. Training loss curve for the proposed model.

Further, the experiments were performed by increasing the number of LSTM layers in the proposed model. It is observed that the performance of the model was declining when the number of layers increased. Table 7 and Fig. 8 present the optimal values of MAPE and R2 with two layers of LSTM.

Table 7. Performance comparison with number of LSTM layers N=1, 2, 3, 4.

Model Type	MAPE (%) ↓	R^2 ↑
Proposed Model (Pre-processing + CNN+LSTM) with 1 Layer (N=1)	3.1250	0.9728
Proposed Model (Pre-processing + CNN+LSTM) with 2 Layers (N=2)	3.0100	0.9743
Proposed Model (Pre-processing + CNN+LSTM) with 3 Layers (N=3)	3.0409	0.9744
Proposed Model (Pre-processing + CNN+LSTM) with 4 Layers (N=4)	3.1449	0.9730

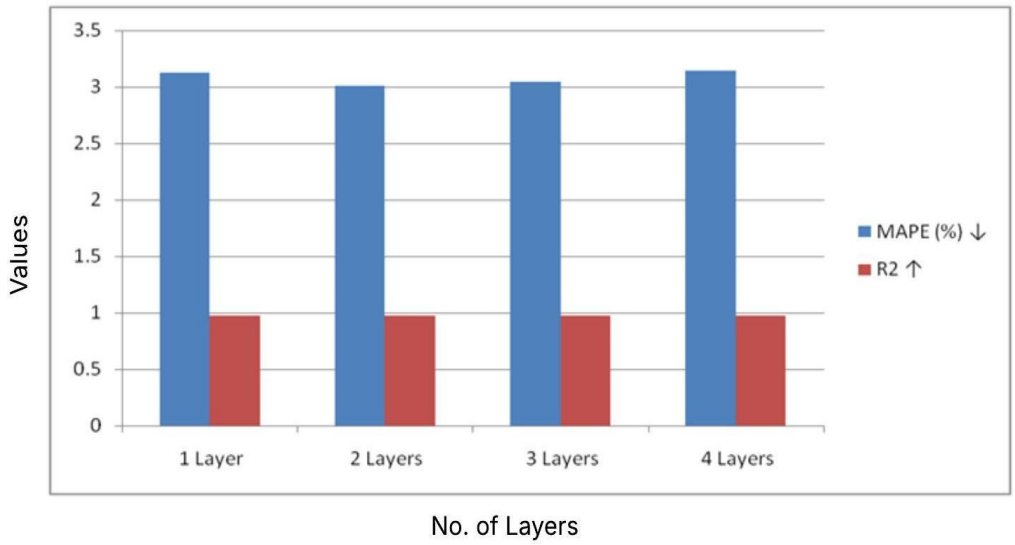


Fig. 8. Performance comparison in terms of MAPE & R2 with number of LSTM layers N=1, 2, 3, 4

4.1. Impact of Time-Context Features (Occupancy Proxies)

To capture variations in energy usage driven by human activity patterns, we integrated a set of temporal context features such as "hour of day", "day of week", "day of month", "month", "week of year", "is_weekend" and is_night, and we used these time-context features to improve our CNN+LSTM model through their addition as engineered temporal and contextual features. These features establish direct measurement of occupancy and act as occupancy proxies. The reason to add these engineered features is:

- Hour of day: Office buildings have peak loads during working hours (9 AM – 6 PM) and reduced loads at night. The short-term load predictions can be improved by capturing this cycle.
- Day of week: Energy consumption patterns differ for weekdays (when offices run full operations) and weekends (when reduced staff and loads occur).
- Is_weekend & is_night: These two are used to explicitly distinguish low-occupancy intervals, so predictions are strong against abrupt drops in demand.
- Month: It is used to record seasonality (e.g., increased HVAC demand in summer or winter).

These exogenous engineered features are crucial for energy usage prediction in any smart educational building because educational building energy consumption for plug-load and lighting operations depends on the operational schedules of such buildings. The proxy features enable the assessment of energy consumption patterns which occur during different times of the day and through various user activities even when direct occupancy monitoring is unavailable. Compared to the suggested CNN+LSTM model without these features, the enhanced model achieved a reduction in RMSE from 0.5610 to 0.5346 (4.7%), MAE from 0.3891 to 0.3713 (4.6%), and MAPE from 3.01% to 2.89% (3.9%), while R^2 increased from 0.9743 to 0.9767. Table 8 and Fig. 9 show the impact of these external occupancy features on the efficiency of the suggested CNN+LSTM model.

These enhancements place significant emphasis on contextual features capable of capturing modern office environments with structured operational cycles wherein energy consumption is dependent on working hours, weekdays–weekends schedules, and inactivity during nighttime hours. Without intrusive sensing, the model proposes a scalable approach to occupant-behavior-aware forecasting by implicitly modeling occupant-driven demand. Practically, this improvement can bring benefits to office building managers who are involved in optimizing HVAC, lighting, and appliance control, while at the same time providing utilities with tools and information that can shape demand response/load management strategies. At a social level, more accurate forecasts save on energy wastage, stabilize grids, and also contribute significantly toward CO_2-free operations of building energy, thus promoting sustainable and efficient energy operations.

Table 8. Impact of engineered features on the accuracy of the proposed model in short-term horizon.

Forecast Horizon	Model Type	RMSE ↓	MAE ↓	MAPE (%) ↓	R^2 ↑
Short-term	Proposed Model (Pre-processing + CNN+LSTM) without engineered features	0.5610	0.3891	3.01	0.9743
	Proposed Model (Pre-processing + CNN+LSTM) with engineered features	**0. 5346**	**0.3713**	**2.89**	**0.9767**

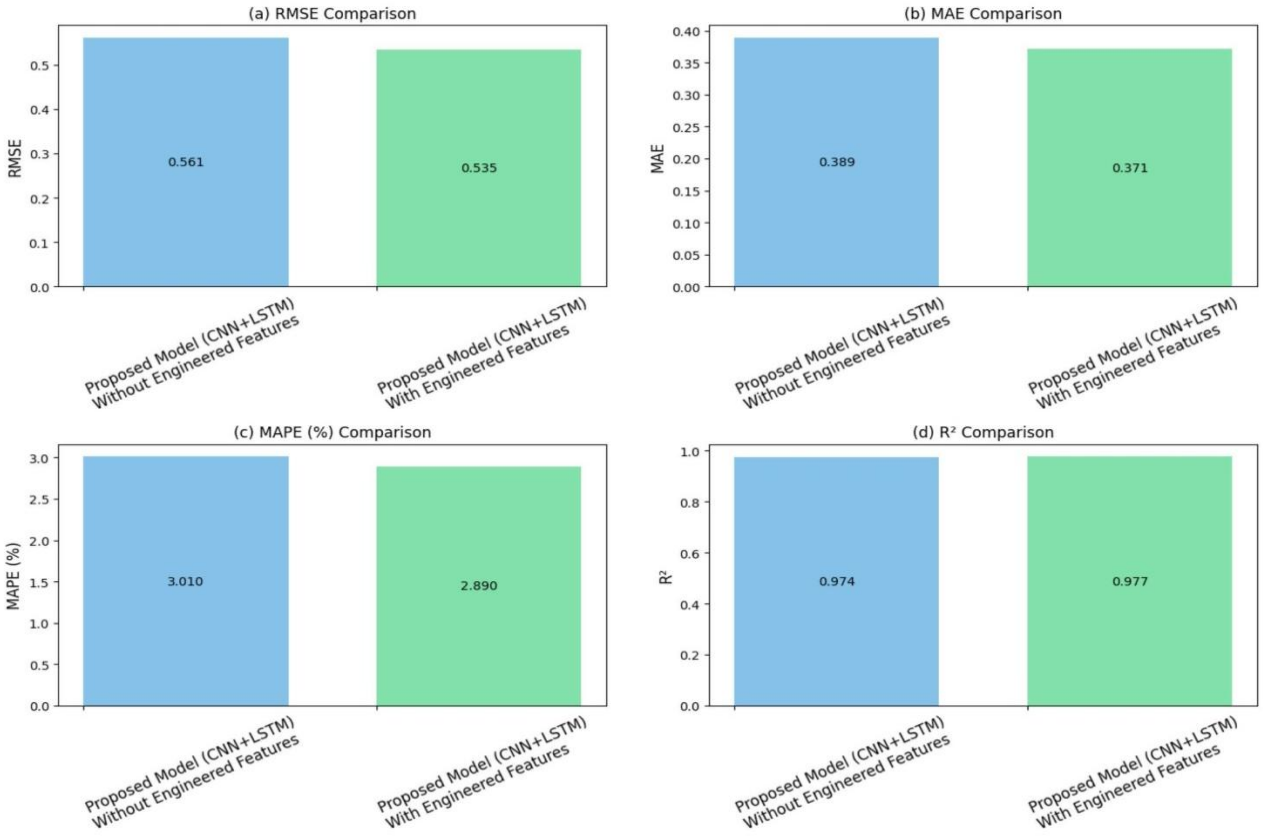


Fig. 9. Short-term performance comparison of proposed model with and without Engineered Features.

The proposed CNN+LSTM hybrid framework with engineered features was also tested on medium-term and long-term horizons and outperformed all other baselines. Table 9 and Fig. 10 present the medium-term forecasting results. For Medium-term forecasting (daily), this framework obtains RMSE of 0.5338 and MAPE of 2.82%, again better than LSTM (MAPE=4.54%), CNN (MAPE=3.60%), and other baselines.

Table 9. Medium-term forecasting comparison of the proposed model.

Model Type	RMSE ↓	MAE ↓	MAPE (%) ↓	R^2 ↑
Pre-processing + XGBoost with engineered features	0.5654	0.3859	3.02	0.9739
Preprocessing + NN with engineered features	0.6222	0.4457	3.68	0.9683
Pre-processing + LSTM with engineered features	0.7629	0.5475	4.54	0.9448
Pre-processing + Stacked LSTM (2 layers) with engineered features	0.5668	0.3935	3.07	0.9738
Pre-processing + Stacked LSTM (3 layers) with engineered features	0.6399	0.4651	3.64	0.9664
Pre-processing + Bi-directional LSTM with engineered features	0.5820	0.4103	3.25	0.9724
Pre-processing + Bi-directional LSTM + Stacked LSTM (2 layers) with engineered features	0.7035	0.5178	4.35	0.9281
Pre-processing + CNN with engineered features	0.6990	0.4876	3.60	0.9589
Pre-processing + TCN with engineered features	0.7540	0.4889	4.79	0.9491
Preprocessing + Transformer with engineered features	3.5036	2.8342	22.91	-0.0002
Proposed Model (Pre-processing + CNN+LSTM) with engineered features	**0.5338**	**0.3626**	**2.82**	**0.9768**

The model accuracy assessment for medium-term forecasting shows the impact of the created features through the results presented in Table 10 and Figure 11. The enhanced model with engineered features shows better performance than the baseline model which lacks those features, because it reduces RMSE from 0.5684 to 0.5338, MAE from 0.3844 to 0.3626, and MAPE from 2.98% to 2.82%. The R^2 value shows a small increase in the range of 0.9736 to 0.9768; the error metrics show better performance because occupancy proxies help detect user patterns that affect energy use.

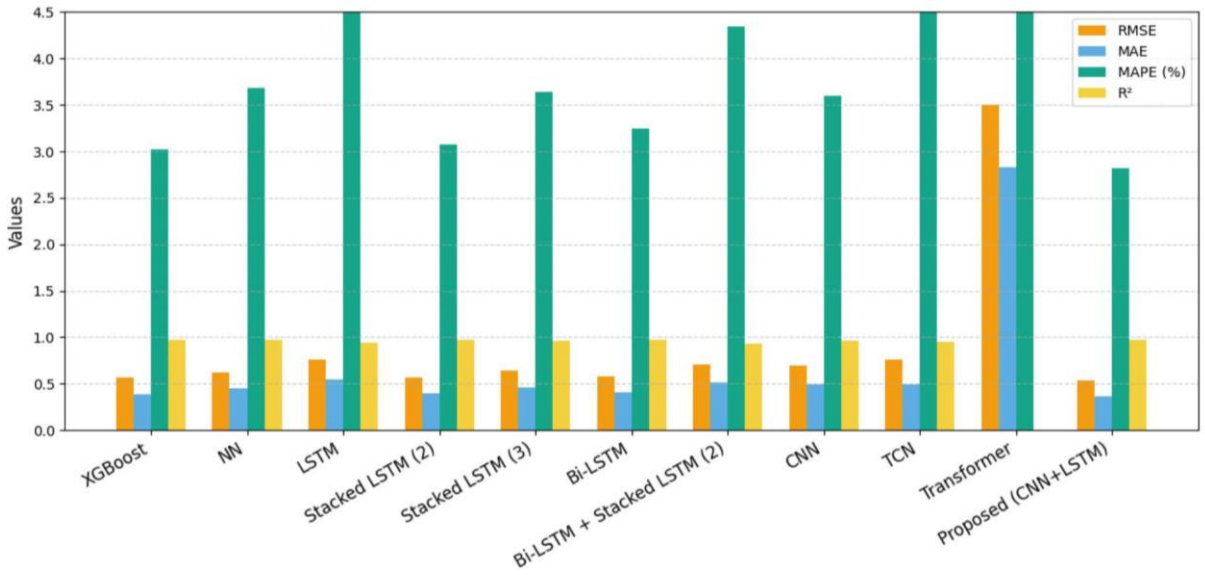


Fig. 10. Medium-term forecasting performance comparison of proposed model (with engineered features) to all the baselines.

Table 10. Impact of engineered features on the accuracy of the proposed model in medium-term horizon.

Forecast Horizon	Model Type	RMSE ↓	MAE ↓	MAPE (%) ↓	R^2 ↑
Medium-term	Proposed Model (Pre-processing + CNN+LSTM) without engineered features-occupancy proxies	0.5684	0.3844	2.98	0.9736
	Proposed Model (Pre-processing + CNN+LSTM) with engineered features-occupancy proxies	**0.5338**	**0.3626**	**2.82**	**0.9768**

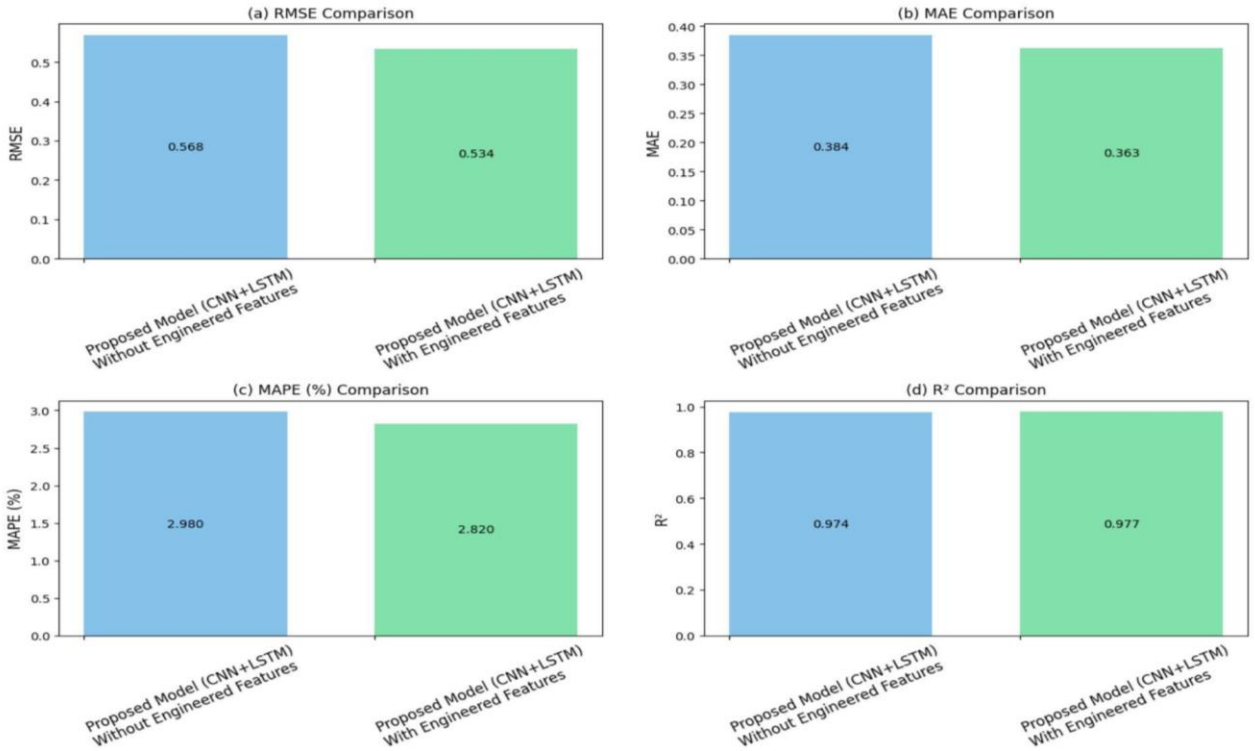


Fig. 11. Medium-term performance comparison of proposed model with and without Engineered Features.

Table 11 and Fig. 12 summarize long-term forecasting performances. Weekly forecasts show the proposed model to be best, achieving the lowest RMSE (0.5464) and MAPE (2.98%) as well as the highest R^2 (0.9757), thus establishing the model's capability to deal with more significant aggregation windows. A constant increase in accuracy across all horizons proves that CNNs extract the local temporal pattern, LSTMs extract the long-term sequential

relationships, and engineered features provide the crux of the actual context; together, these form a generalizable architecture that aids a gamut of real-world applications from demand response and operational scheduling to long-term energy planning and sustainable grid management.

Table 11. Long-term forecasting comparison of the proposed model with all the baselines.

Model Type	RMSE ↓	MAE ↓	MAPE (%) ↓	R^2 ↑
Pre-processing + XGBoost with engineered features	0.5674	0.3890	3.05	0.9738
Preprocessing + NN with engineered features	0.5908	0.4390	3.75	0.9714
Pre-processing + LSTM with engineered features	0.6534	0.4480	3.51	0.9633
Pre-processing + Stacked LSTM (2 layers) with engineered features	0.5532	0.3865	3.02	0.9750
Pre-processing + Stacked LSTM (3 layers) with engineered features	0.5682	0.3917	3.04	0.9737
Pre-processing + Bi-directional LSTM with engineered features	0.5643	0.3927	3.12	0.9740
Pre-processing + Bi-directional LSTM + Stacked LSTM (2 layers) with engineered features	0.6855	0.5059	4.25	0.9317
Pre-processing + CNN with engineered features	0.7332	0.5197	3.90	0.9561
Pre-processing + TCN with engineered features	0.7685	0.5144	4.93	0.9477
Preprocessing + Transformer with engineered features	3.5035	2.8306	22.87	-0.0001
Proposed Model (Pre-processing + CNN+LSTM) with engineered features	**0.5464**	**0.3780**	**2.98**	**0.9757**

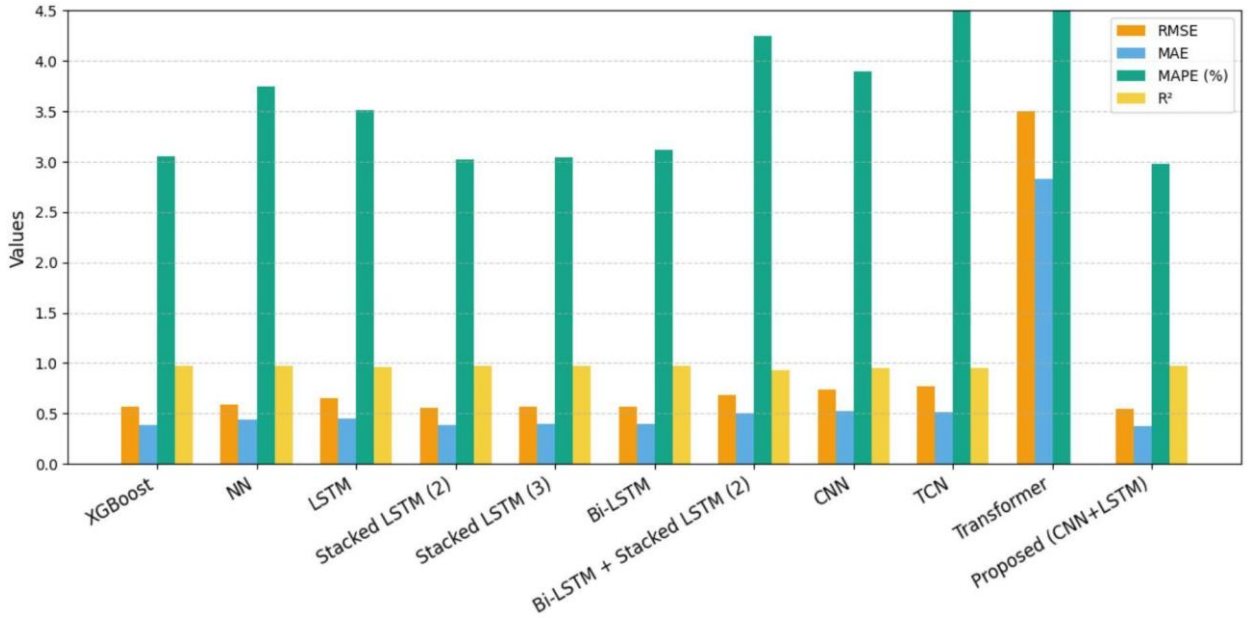


Fig. 12. Long-term forecasting performance comparison of proposed model with engineered features to all the baselines with engineered features.

The model's results in long-term forecasting, which show the impact of the engineered features, are presented in Table 12 and Figure 13. The long-term forecasting accuracy improves when engineered features are included in the forecasting model. The enhanced model produces better results because it achieves a lower RMSE of 0.5464 compared to its previous value of 0.5609 and it shows an improved R^2 value of 0.9757 which exceeds its former R^2 value of 0.9743. The MAE and MAPE values show very small changes because both models maintain consistent error patterns throughout their operation.

Table 12. Impact of engineered features on the accuracy of the proposed model in long-term horizon.

Forecast Horizon	Model Type	RMSE ↓	MAE ↓	MAPE (%) ↓	R^2 ↑
Long-term	Proposed Model (Pre-processing + CNN+LSTM) without engineered features-occupancy proxies	0.5609	0.3734	2.99	0.9743
	Proposed Model (Pre-processing + CNN+LSTM) with engineered features-occupancy proxies	**0.5464**	**0.3780**	**2.98**	**0.9757**

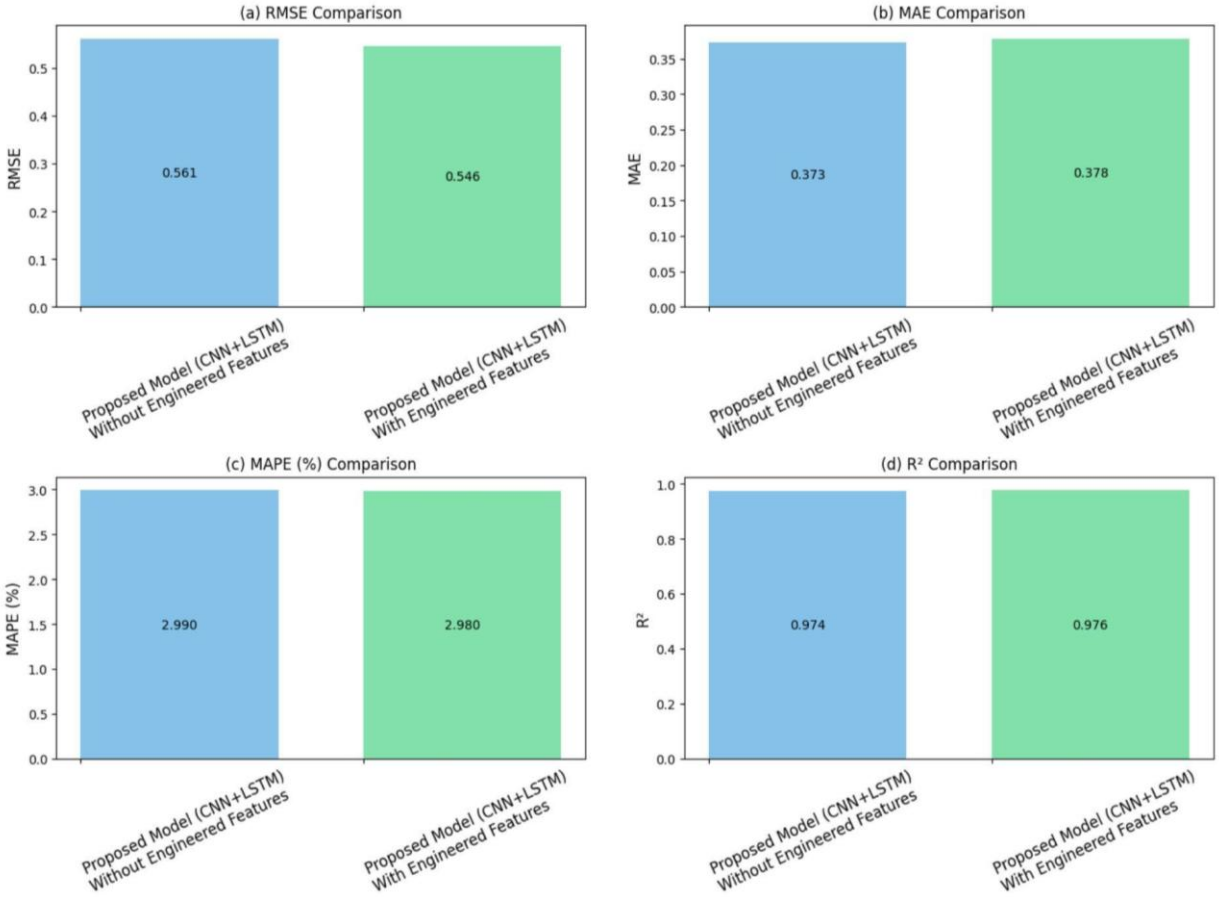


Fig. 13. Long-term performance comparison of proposed model with and without Engineered Features.

4.2. Statistical Significance Testing: Paired t-Test & Diebold-Mariano Test

To determine whether the model's enhancements through engineered temporal features lead to statistically significant results, we conducted a paired t-test and DM test to evaluate forecast errors from two competing models, one that used engineered features and one that did not, across all five folds of testing. The results shown in Table 13 demonstrate that the engineered-feature model achieved a prediction error reduction that reached highly significant statistical levels with a p-value less than 0.05. The results show that the RMSE/MAE/MAPE improvements do not exist because of random variations, but instead provide consistently better performance across different testing folds.

Table 13. Results of paired t-test over the forecast errors of the CNN+LSTM model (with vs. without engineered features) across all 5 folds.

S. No.	Test Name	Forecast Horizon	T-Statistic	P-value	Improvement with engineered features
1	Paired t-Test	Short-term	25.8175	< 0.000001	Statistically significant (p < 0.05)
		Medium-term	11.9835	< 0.000001	Statistically significant (p < 0.05)
		Long-term	-44.3299	< 0.000001	Statistically significant (p < 0.05)
2	Diebold-Mariano Test	**Forecast Horizon**	**DM-Statistic**	**P-value**	**Improvement with engineered features**
		Short-term	-8.6536	< 0.000001	Statistically significant (p < 0.05)
		Medium-term	-9.4275	< 0.000001	Statistically significant (p < 0.05)
		Long-term	-6.5544	< 0.000001	Statistically significant (p < 0.05)

The paired test was performed across all three horizons—short, medium, and long term—to examine whether the model's improvements are statistically significant or not. The results of this test confirm that the improvements achieved by engineered features in the proposed model are statistically significant. The p-value in short-term, medium-term, and long-term is less than 0.05 ($p<0.000001$), which proves that the decline in forecast error rate is highly significant.

Likewise, the Diebold–Mariano test performed across all three horizons yielded DM statistics of -8.6536, -9.4275, and -6.5544 with a p-value < 0.000001, indicating a statistically significant difference in forecast accuracy between the two models. The (-DM) value shows that the model with engineered features provides good results as compared to the model without engineered features. The enhancement in the model's forecasting accuracy can be considered highly statistically significant (as $p<0.05$), indicating that the engineered features also meaningfully improve performance in the long-term horizon.

4.3. Limitations

This study has several limitations that should be acknowledged. The dataset only includes three months of data which prevents researchers from studying seasonal patterns and extended behavioral changes. The analysis relies on data from one research building, which lacks the equipment and operational patterns of other academic structures, student housing, and smart-campus systems. The study used engineered features as time-context proxies because they lack direct occupancy measurement which creates a restriction on model performance used to study energy consumption patterns associated with human activities. The study establishes important device-level energy forecasting methods for research-focused smart-campus buildings while building a future research base through its implementation of extended data collection across multiple facilities with advanced sensor systems.

5. Conclusion & Future Scope

The work describes a CNN-LSTM-based hybrid deep learning model proposed for forecasting building energy consumption across short-, medium-, and long-term horizons. The framework used CNNs for modeling local patterns and LSTMs for modeling sequential dependencies and produced substantial improvements compared to some baseline models. The temporal engineered features further enhanced accuracy and interpretability, showing that even simple and very parsimonious time-context variables can greatly improve predictive capability. The results demonstrate that the proposed method with engineered features performs better in a consistent and statistically significant manner compared to baseline models: For short-term forecasting, the improvements were 4.71% and 3.99% in RMSE and MAPE, respectively; for medium-term, 6.08% and 5.37%, respectively; and for long-term, 2.58% and 0.33%, respectively. Although the architecture itself is not novel, the domain-specific adaptation, feature engineering strategy, and unified multi-horizon design contribute new insights into device-level load forecasting for smart campuses. Therefore, building up an efficient and interpretable forecasting tool like ours can build a framework for suppliers, policymakers, and building managers to engage in demand-side management, grid-aware scheduling, and sustainable energy system-oriented operations in modern academic infrastructures. Future work will incorporate longer datasets across full seasonal cycles and explore true occupancy measurements or sensor fusion to further enhance behavioral modeling and long-term prediction robustness.

All the Declarations and Statements

Author Contributions Statement

Ajay Kumar – Writing – Drafted the initial manuscript, Writing – Review and Editing, contributed to the literature survey, Conceptualization, Methodology, Formal Analysis, Model Training, Validation, and Performance Evaluation: Led the model training process, validated results using standard metrics, and benchmarked performance against existing methods, Formal Analysis, Visualization, and Statistical Analysis: Performed in-depth analysis of experimental results, prepared performance charts, and ensured the statistical robustness of the evaluation.
Rainu Nandal – Writing – Review and Editing, and Project Management: Reviewed and edited the manuscript, ensured clarity and coherence, and helped coordinate project milestones and deadlines, and Supervision: Proposed research ideas, Constructed the overall framework, and supervised project execution.
Kamaldeep Joshi – Data Curation and Software Implementation: Handled data acquisition, dataset preprocessing, and implementing the research model, Writing – Review and Editing.

All authors have read and agreed to the published version of the manuscript.

Conflict of Interest Statement

The authors declare no conflicts of interest.

Funding Declaration

None.

Data Availability Statement

This study analyzed publicly available datasets. The results can be obtained on request to the corresponding author. The datasets can be found here: “https://data.mendeley.com/datasets/g392vt7db9/1”, accessed on “22 July 2025”.

Ethical Declarations

This study does not involve human participants or animals. The dataset used consists solely of building-level energy consumption measurements. Therefore, ethical approval was not required for this research.

Acknowledgments
We sincerely thank the experts for their professional evaluation and valuable recommendations, which have contributed to improving the quality of the experiment and the reliability of its results.

Declaration of Generative AI in Scholarly Writing
The authors used a generative AI tool solely to strengthen the language and readability of the manuscript through language editing. No AI-generated text, analysis, data interpretation, or scientific content was used. The authors take full responsibility for the content of this work.

Abbreviations
The following abbreviations are used in this manuscript:

AI - Artificial Intelligence
ANN - Artificial Neural Networks
ARMAX - Autoregressive Moving Average with Exogenous
ASHRAE - American Society of Heating, Refrigerating, and Air Conditioning
BiLSTM - Bidirectional LSTM
CEEMD - Complementary Ensemble Empirical Mode Decomposition
CNN - Convolutional Neural Networks
DEMD - Differential Empirical Mode Decomposition
DL - Deep Learning
DM Test - Diebold–Mariano Test
DNN - Deep Neural Networks
DOE - Department of Energy
EMD - Empirical Mode Decomposition
GA - Genetic Algorithm
IQR - Interquartile Range
KF - K-Fold Cross Validation
LSTM - Long Short-Term Memory Networks
MAE - Mean Absolute Error
MAPE - Mean Absolute Percentage Error
MetaFA-LSSVR - Metaheuristic Firefly Algorithm with Least Squares Support Vector Regression
MSE - mean squared error
NLP - Natural Language Processing
NREL - National Renewable Energy Laboratory
N-BEATS - Neural Basis Expansion Analysis for Time Series
PSO - Particle Swarm Optimization
R^2- Coefficient of Determination
RF - Random Forest
RMSE - Root Mean Square Error
RNN - Recurrent Neural Networks
RSF - Research Support Facility
SARIMA - Seasonal Autoregressive Integrated Moving Average
SGs - Smart Grids
SSA - Sparrow Search Algorithm
SVM - Support Vector Machines
XGBoost - eXtreme Gradient Boosting

References

[1] M. Q. Raza, and A. Khosravi, “A review on artificial intelligence-based load demand forecasting techniques for smart grid and buildings,” *Renew. Sustain. Energy Rev*., vol. 50, pp. 1352–1372, 2015. doi:10.1016/j.rser.2015.04.065.
[2] N. Wei, C. Li, X. Peng, F. Zeng, X. Lu, “Conventional models and artificial intelligence-based models for energy consumption forecasting: a review,” *J. Petroleum Sci. Eng*., vol. 181, pp. 106187, 2019. doi:10.1016/j.petrol.2019.106187.
[3] D. Kaur, S. N. Islam, M. Mahmud, Z. Dong, “Energy forecasting in smart grid systems: a review of the state-of-the-art techniques,” *arXiv Prepr*, 2020. arXiv:2011.12598.
[4] L. Pérez-Lombard, J. Ortiz, and C. Pout, “A review on buildings energy consumption information,” *Energy Build*., vol. 40, no. 3, pp. 394–398, 2008. doi:10.1016/j.enbuild.2007.03.007.

[5] Y. Ji, K.J. Lomas, and M.J. Cook, "Hybrid ventilation for low energy building design in south China," *Build. Environ.*, vol. 44, pp. 2245–2255, 2009. doi:10.1016/j.buildenv.2009.02.015.
[6] S. Manu, Y. Shukla, R. Rawal, L.E. Thomas, and R. De Dear, "Field studies of thermal comfort across multiple climate zones for the subcontinent India Model for Adaptive Comfort (IMAC)," *Build. Environ.*, vol. 98, pp. 55–70, 2016. doi:10.1016/j.buildenv.2015.12.019.
[7] E. Delzendeh, S. Wu, A. Lee, and Y. Zhou, "The impact of occupants' behaviours on building energy analysis: A research review," *Renew. Sustain. Energy Rev.*, vol. 80, pp. 1061–1071, 2017. doi:10.1016/j.rser.2017.05.264.
[8] M. I. Ahmad, "Seasonal decomposition of electricity consumption data," *Rev. Integr. Bus. Econ. Res*., vol. 6, no. 4, pp. 271, 2017
[9] H. Kaur, and S. Ahuja, "Time series analysis and prediction of electricity consumption of health care institution using ARIMA model," *In Proc. Sixth Int. Conf. Soft Comput. Probl. Solving*, Singapore, pp. 347–358, 2017, December 2016. doi:10.1007/978-981-10-3325-4_35.
[10] K. Amarasinghe, D. L. Marino, and M. Manic, "Deep neural networks for energy load forecasting," *In 2017 IEEE 26th Int. Symp. Ind. Electron. (ISIE), Edinburgh, United Kingdom, pp. 1483–1488*, June 2017. doi:10.1109/ISIE.2017.8001465.
[11] S. Nazir, A. Ab Aziz, J. Hosen, N. A. Aziz, G. R. Murthy, "Forecast energy consumption time-series dataset using multistep LSTM models," *J. Phys. Conf. Ser*., vol.1933, no. 1, pp. 012054, 2021. doi:10.1088/1742-6596/1933/1/012054.
[12] J.G. Jetcheva, M. Majidpour, W.P. Chen, "Neural network model ensembles for building-level electricity load forecasts," *Energy Build.*, vol. 84, pp. 214–223, 2014. doi:10.1016/j.enbuild.2014.08.004
[13] M.A. Khairalla, X. Ning, N.T. AL-Jallad, M.O. El-Faroug, "Short-Term Forecasting for Energy Consumption through Stacking Heterogeneous Ensemble Learning Model," *Energies*, vol. 11, pp. 1605, 2018. doi:10.3390/en11061605.
[14] J. Geng, , M.L. Huang, M.W. Li, W.C. Hong, "Hybridization of seasonal chaotic cloud simulated annealing algorithm in a SVR-based load forecasting model," *Neurocomputing*, vol. 151, pp. 1362–1373, 2015. doi:10.1016/j.neucom.2014.10.055.
[15] G.F. Fan, L.L. Peng, W.C. Hong, and F. Sun, "Electric load forecasting by the SVR model with differential empirical mode decomposition and auto regression," *Neurocomputing*, vol. 173, pp. 958–970, 2016. doi:10.1016/j.neucom.2015.08.051.
[16] Z. Wang, and R.S. Srinivasan, "A review of artificial intelligence based building energy prediction with a focus on ensemble prediction models," *In 2015 Winter simul. conf. (WSC)*, pp. 3438–3448, 2016. doi:10.1109/WSC.2015.7408504.
[17] P.W. Tien, S. Wei, J. Darkwa, et al, "Machine learning and deep learning methods for enhancing building energy efficiency and indoor environmental quality - a review," *Energy AI*, vol. 10, pp. 100198, 2022. doi:10.1016/j.egyai.2022.100198.
[18] J. Lago, G. Marcjasz, B. De Schutter et al, "Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark," *Appl. Energy*, vol. 293, no. 116, pp. 983, 2021. doi:10.1016/j.apenergy.2021.116983.
[19] J.S. Chou, and N.T. Ngo, "Time series analytics using sliding window metaheuristic optimization-based machine learning system for identifying building energy consumption patterns," *Appl. Energy*, vol. 177, pp. 751–770, 2016. doi:10.1016/j.apenergy.2016.05.074.
[20] Y. Kim, H.G. Son, and S. Kim, "Short term electricity load forecasting for institutional buildings," *Energy Rep*., vol. 5, pp. 1270–1280, 2019. doi:10.1016/j.egyr.2019.08.086.
[21] T. Pinto, I. Praça, Z. Vale, and J. Silva, "Ensemble learning for electricity consumption forecasting in office buildings," *Neurocomputing*, vol. 423, pp. 747–755, 2021. doi:10.1016/j.neucom.2020.02.124.
[22] G. Chitalia, M. Pipattanasomporn, V. Garg, and S. Rahman, "Robust short-term electrical load forecasting framework for commercial buildings using deep recurrent neural networks," *Appl. Energy*, vol. 278, pp. 115410, 2020. doi:10.1016/j.apenergy.2020.115410.
[23] F. Pallonetto, C. Jin, and E. Mangina, "Forecast electricity demand in commercial building with machine learning models to enable demand response programs," *Energy AI*, vol. 7, pp. 100121, 2022. doi:10.1016/j.egyai.2021.100121.
[24] J. Hwang, D. Suh, and M.O. Otto, "Forecasting electricity consumption in commercial buildings using a machine learning approach," *Energies*, 13, pp. 5885, 2020. doi:10.3390/en13225885.
[25] X. Wen, J. Liao, Q. Niu, N. Shen, and Y. Bao, "Deep learning-driven hybrid model for short-term load forecasting and smart grid information management", *Sci. Rep*., vol. 14, pp. 13720, 2024. doi:10.1038/s41598-024-63262-x.
[26] H. Li, S. Li, Y. Wu, Y. Xiao, Z. Pan, and M. Liu, "Short-term power load forecasting for integrated energy system based on a residual and attentive LSTM-TCN hybrid network". *Front. Energy Res*., vol. 12, pp. 1384142, 2024. doi:10.3389/fenrg.2024.1384142.
[27] J. Bai, W. Zhu, S. Liu, C. Ye, P. Zheng, and X. Wang, "A Temporal Convolutional Network–Bidirectional Long Short-Term Memory (TCN-BiLSTM) Prediction Model for Temporal Faults in Industrial Equipment", Appl. Sci., vol.15, no. 4, p. 1702, 2025. https://doi.org/10.3390/app15041702
[28] B. Parizad, H. Ranjbarzade, A. Jamali, and H. Khayyam, "An Intelligent Hybrid Machine Learning Model for Sustainable Forecasting of Home Energy Demand and Electricity Price", *Sustainability*, vol. 16, no. 6, pp. 2328, 2024. doi:10.3390/su16062328.
[29] Z. Bai, "Residential electricity prediction based on GA-LSTM modeling", *Energy Rep*., vol. 11, pp. 6223-6232, 2024. doi:10.1016/j.egyr.2024.06.010.
[30] Y. Özüpak, and S. Mansurov, "Optimizing electricity demand forecasting with a novel RNN-LSTM hybrid model", *Energy Sources B: Econ. Plan. Policy*, vol. 20, no. 1, 2025. doi:10.1080/15567249.2025.2531448.
[31] D. Mariano-Hernández, et al., "Comparative study of continuous hourly energy consumption forecasting strategies with small data sets to support demand management decisions in buildings," *Energy Sci. Eng.*, vol. 10, pp. 4694–4707, 2022. doi:10.1002/ese3.1298.
[32] R. Spencer, S. Ranathunga, M. Boulic, A.H. van Heerden, and T. Susnjak, "Transfer learning on transformers for building energy consumption forecasting—A comparative study". *Energy Build.*, vol. 336, pp. 115632, 2025. doi:10.1016/j.enbuild.2025.115632.

[33] D. Fernández-Martínez, and M.A. Jaramillo-Morán, "Multi-Step Hourly Power Consumption Forecasting in a Healthcare Building with Recurrent Neural Networks and Empirical Mode Decomposition," *Sensors*, vol. 22, pp. 3664, 2022. doi:10.3390/s22103664.
[34] E. Skomski, J.Y. Lee, W. Kim, Chandan, S. Katipamula, and B. Hutchinson, "Sequence-to-sequence neural networks for short-term electrical load forecasting in commercial office buildings," *Energy Build*., vol. 226, pp. 110350, 2020. doi:10.1016/j.enbuild.2020.110350.
[35] I. Karijadi, and S.Y. Chou, "A hybrid RF-LSTM based on CEEMDAN for improving the accuracy of building energy consumption prediction," *Energy Build*., vol. 259, pp. 111908, 2022. doi:10.1016/j.enbuild.2022.111908.
[36] L. Lei, S.Shao, L. Liang, "An evolutionary deep learning model based on EWKM, random forest algorithm, SSA, and BiLSTM for building energy consumption prediction," *Energy*, vol. 288, pp. 129795, 2024. doi:10.1016/j.energy.2023.129795.
[37] Z. Ding, W. Chen, T. Hu, X. Xu, "Evolutionary double attention-based long short-term memory model for building energy prediction: Case study of a green building," *Appl. Energy*., vol. 288, pp. 116660, 2021. doi:10.1016/j.apenergy.2021.116660.
[38] B. Doherty, and K. Trenbath, "Data for: Device-level plug load disaggregation in a zero energy office building and opportunities for energy savings," *Mendeley Data*., vol. 06, 2019, doi: 10.17632/g392vt7db9.1.
[39] J. Sessa, and D. Syed, "Techniques to deal with missing data," *In Proc. 5th Int. Conf. Electron. Devices, Syst. Appl. (ICEDSA)*, Sarawak, Malaysia. pp. 1–4, December 2016. doi: 10.1109/ICEDSA.2016.7818486.
[40] S. V. Khond, "Effect of data normalization on accuracy and error of fault classification for an electrical distribution system," *Smart Sci*., vol. 8, no. 3, pp. 117–124, 2020. doi: 10.1080/23080477.2020.1799135.
[41] S. Bai, V. Koltun, and J.Z. Kolter, "Multiscale deep equilibrium models." *NIPS'20: Proc. 34th Int. Conf. Neural Inf. Process. Syst..*, Red Hook, United States, vol. 33, pp. 5238-5250, December 2020.
[42] S. Kasmaiee, S. Kasmaiee, and M. Homayounpour, "Correcting spelling mistakes in Persian texts with rules and deep learning methods," *Sci. rep.*, vol. 13, no. 1, pp. 19945, 2023. doi:10.1038/s41598-023-47295-2.
[43] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," Journal of Machine Learning Research, vol. 15, pp. 1929–1958, 2014.
[44] B. G. Marcot, and Hanea, A. M., "What is an optimal value of k in k-fold cross-validation in discrete Bayesian network analysis?," *Comput. Statis*t., vol. 36, no. 3, pp. 2009–2031, 2021. doi: 10.1007/s00180-020-00999-9.
[45] B. Devanathan, P. Manelllore, S.A. Lakshmanan, S. Kayalvizhi,, and A. Suyampulingam, "AI-Based Energy Consumption Forecasting in Smart Homes Using Multi-Algorithm Analysis Across Temporal Resolutions, *Results Eng.*, vol. 29, pp. 109272. 2026. doi:10.1016/j.rineng.2026.109272.
[46] B.K. Hammerschmitt, M.V.H. Rambo, A.D.S. Leone, L.M. Iantorno, H.B. Schiavon, D.P. Corrêa, P. Lissa, M. Keane, and R.J. Riella, "Deep Learning for Residential Electrical Energy Consumption Forecasting: A Hybrid Framework with Multiscale Temporal Analysis and Weather Integration", *Energies*, vol. 18, no. 22, pp. 5885, 2025. doi:10.3390/en18225885.
[47] D. Syed, S. S. Refaat, H. Abu-Rub, "Performance evaluation of distributed machine learning for load forecasting in smart grids," *In Proc. Cybern. Informat. (K&I).*, pp. 1–6, 2020. doi: 10.1109/KI48306.2020.9039797.

Authors' Profiles



Ajay Kumar, Research Scholar (Ph.D.) in the Department of Computer Science and Engineering at University Institute of Engineering and Technology, Maharshi Dayanand University, Rohtak, India. He has received his M. Tech (Computer Engineering) from J.C. Bose University of Science and Technology, Faridabad, India (formerly YMCA University of Science and Technology) and B. Tech (Computer Engineering) from University Institute of Engineering and Technology, M.D.U., Rohtak. He has more than 12 years of academic experience. He has contributed several research papers in international journals, international/national conferences/proceedings and edited books. His areas of research are machine learning, IoT, and network security. He has organized various workshops and FDPs in these areas. He is the lifetime member of International Association of Engineers (IAENG).



Rainu Nandal, Ph.D., is an Associate Professor in the Department of Computer Science and Engineering at University Institute of Engineering and Technology, Maharshi Dayanand University, Rohtak, India. She received her Ph.D. degree in Computer Science and Engineering from Department of Computer Science and Engineering at UIET, Maharshi Dayanand University, Rohtak, India, 2017. She has more than 18 years of academic experience. She is the author of more than 72 scientific contributions including articles in international Peer-review Journals in the area of Machine Learning and Computer Vision. Her area of interest includes Machine Learning, Big Data, Informatics and Control Theory.



Kamaldeep Joshi, Ph.D., is an Assistant Professor in the Department of Computer Science and Engineering at University Institute of Engineering and Technology, Maharshi Dayanand University, Rohtak, India. He received his Ph.D. degree in Computer Science and Engineering from Department of Computer Science and Engineering at UIET, Maharshi Dayanand University, Rohtak, India, 2017. He has more than 12 years of academic experience. He is the author of more than 63 scientific contributions including articles in international Peer-review Journals and conferences in the area of Machine Learning and Computer Vision. He has completed two minor research projects and has been granted one Indian patent. His area of interest includes Machine Learning, Image Processing and Steganography.

How to cite this paper

Ajay Kumar, Rainu Nandal, Kamaldeep Joshi, "Multi-Horizon Deep Learning Framework for Sustainable Energy Forecasting in Smart Campuses: Enabling Energy Efficient Smart Educational Buildings", International Journal of Modern Education and Computer Science(IJMECS), Vol.18, No.5, pp. 168-191, 2026. DOI:10.5815/ijmecs.2026.05.10